ECOLOGY AND EPIDEMIOLOGY OF *SALMONELLA* AND *LISTERIA MONOCYTOGENES* IN NEW YORK STATE PRODUCE PRODUCTION ENVIRONMENTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Laura Kathryn Strawn

January 2014

ECOLOGY AND EPIDEMIOLOGY OF *SALMONELLA* AND *LISTERIA MONOCYTOGENES* IN NEW YORK STATE PRODUCE PRODUCTION ENVIRONMENTS

Laura Kathryn Strawn, Ph.D.

Cornell University 2014

Foodborne pathogen contamination of produce in the preharvest environment continues to present a considerable challenge and can lead to recalls or even outbreaks. Thus, there is a need for further development of science-based approaches to assist growers in minimizing the risk of preharvest contamination of produce. *Salmonella* and *Listeria monocytogenes* are two foodborne pathogens of concern in produce that both cause severe illness in humans. In these studies, we employed several approaches to model *Salmonella* and *L. monocytogenes* contamination in the preharvest environment as an ecological process. Specifically, the studies presented here investigated (i) the use of Geographical Information Systems (GIS) modeling to predict pathogen prevalence based on landscape and meteorological factors, (ii) the association between pathogen presence and field management practices to quantify likelihood of preharvest contamination, and (iii) the application of subtyping-based source tracking of pathogens in the preharvest environment. We observed that *Salmonella* and *L. monocytogenes* prevalence was not uniform across croplands, and that prevalence in a specific area may be significantly higher based on specific factors. For example, the localized prevalence of *L. monocytogenes* was higher in areas close to water. We also showed that *Salmonella* and *L. monocytogenes* isolation in fields was significantly influenced by specific management factors and that many of those factors were time dependent. For instance, application of manure to a field within a year increased the odds (OR=16.7, 95% CI=3.0, 94.4) of a *Salmonella*-positive field. Lastly, we found that *Salmonella* serotypes were diverse in the preharvest environment, and that PFGE type varied based on geographic region. Collectively, these findings suggest that risk management of pathogen

contamination in the preharvest environment requires tailoring to each specific farm, as each farm has a unique set of factors that influence the risk of contamination. Knowledge of such factors for a farm will allow growers to prioritize risks. The ability to prioritize risks on produce farms will develop a preventative approach to preharvest food safety. Produce growers will be able to target areas within their farms that are at high risk for contamination, or implement more informed field management practices to reduce potential contamination of produce.

BIOGRAPHICAL SKETCH

Laura Kathryn Strawn was born on October 5, 1984 in Long Beach, California as the oldest of three children to John and Kathy Strawn. Laura attended the University of California, Davis from 2003-2007 where she graduated with a B.S. degree in Food Science with an emphasis in Microbiology. As an undergraduate student, Laura was a part time technician in a food microbiology laboratory under the direction of Dr. Linda Harris where she discovered her passion for food safety. Laura decided to pursue a M.S. degree under the direction of Dr. Michelle Danyluk at the University of Florida. Following completion of her degree in July 2009, Laura began a Ph.D. program in Food Science, Epidemiology and Microbiology under the direction of Drs. Martin Wiedmann, Randy Worobo, and Yrjo Gröhn in August 2009. Laura will begin a position as a post-doctoral researcher in Dr. Wiedmann's research group in January 2014. Outside of her food microbiology life, Laura enjoys traveling, watching sports, and drinking wine with friends.

Dedicated To My Family

Dad, Mom, Matt, Zach and Salem

ACKNOWLEDGMENTS

I am so fortunate to be able to call Dr. Martin Wiedmann my advisor. I would not be the scientist and person I am today without him in my life. Martin has always pushed me to strive for "a higher bar" in my graduate studies. He has provided me with an abundance of opportunities to prepare and assist me in my future career. I am grateful for his constant support, even when things get a bit tough. Lastly, for his open door policy- Martin always made time to discuss research ideas with me; in addition, to providing advice and guidance when I needed it.

I would also like to thank my committee members, Dr. Randy Worobo and Dr. Yrjo Gröhn. I am eternally grateful to Randy for his constant support and willingness to discuss my research ideas and thoughts. Randy taught me how to communicate with a produce grower, as well as provided me with the opportunity to work on an amazing produce safety grant. I appreciate the perspective Yrjo has added to my research, as well as the time and support he has provided in discussing my research.

I am particular grateful to Dr. Betsy Bihn for her thoughtful discussion of my research projects; in addition to her friendship. I would also like to thank her for stressing the importance of the applied nature of my work for the produce grower. I would also like to thank Dr. Kathryn Boor for being a wonderful role model. I strive to have Kathryn's professionalism and class; she is truly a special person. Lastly, I would like to thank Dr. Alicia Orta-Ramirez for her constant support throughout my graduate studies.

I cannot go without thanking everyone at the Food Safety Laboratory, both past and present. Everyone at the FSL has become my East Coast family. I could not have completed my graduate studies without their support. Special thanks to all the technicians in the FSL: Maureen Gunderson, Sherry Roof, Barbara Bowen, Steve Warchocki, Lexi Andrus, Shelly Clark, Dave Kent, Emily Wright, and Nicole Martin.

I would also like to thank my family and friends for their support throughout my graduate studies. I appreciate all the time and emotional support they provided me, as well as all the sacrifices they have made. Particular thanks to Esther Fortes and Dr. Marianne Fatica, I love you both dearly. Finally I would like to thank those who have been supportive of me along the way, including Kathryn Sullivan, Stephanie Masiello Schuette, Daina Ringus, Matt Stasiewicz, Travis Chapin, Kitiya Vongkamjam, and Lorraine Rodriguez-Rivera.

This work was supported by an AFRI Grant No. 2012-67011-19875 (L.K. Strawn and M. Wiedmann) from the USDA NIFA. Additionally, this research was supported by USDA NIFSI-2008-51110-0688 (R.W. Worobo) and NIFSI-2008-51110-04333 (K.N. Nightingale and M. Wiedmann).

Lastly, I would like to thank all the New York State produce growers who participated in my research studies for your time and cooperation. You were invaluable to the success of this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Salmonella and *Listeria monocytogenes* are two foodborne pathogens of significant concern to public health, as each pathogen accounts for an estimated 378 and 255 deaths annually in the United States (US), respectively (1). Some of these deaths are a result of consumption of contaminated fresh fruits and vegetables (1, 2). Contamination of produce can occur at virtually any point along the farm to fork continuum; however, several of the largest produce-borne outbreaks have been traced back to the production environment. In 2008, an outbreak of *Salmonella* Saintpaul in jalapeno peppers sickened 1,500 individuals across 43 US states and Canada. The trace back investigation determined *S*. Saintpaul was isolated from peppers in a field and an agricultural water source used for irrigation from one of the farms implicated in the outbreak (3). Additionally, in 2012, a cantaloupe-borne *Salmonella (*serotypes Typhimurium and Newport) outbreak caused 261 illnesses and 3 deaths in 24 US states. Upon investigation, it was determined the source of the outbreak was one farm, with initial contamination of the cantaloupes occurring in the production environment (4). While no *L. monocytogenes* outbreaks have been traced back to the preharvest environment, the risk of produce contaminated with *L. monocytogenes* is serious. In 2011, there were 147 illnesses, 33 deaths, and 1 miscarriage due to a *L. monocytogenes* outbreak in cantaloupe (5). *L. monocytogenes* presents unique challenges for the produce industry; as it can be found in diverse environments, often at considerable prevalence, and is able to persist and amplify in refrigeration conditions (6, 7). One study (8) demonstrated that *L. monocytogenes* was able to grow at faster rates under produce storage conditions (7-15°C), compared with *Salmonella*. Thus, preventative strategies focused on reducing potential *L. monocytogenes* contamination in the field would be

fundamental to decreasing the risks of postharvest contamination of produce by *L. monocytogenes*. Collectively, the role of the environment on foodborne disease outbreaks may be best understood by field studies on commercial produce farms that focus on foodborne pathogen distribution, prevalence, persistence and diversity in the preharvest environment.

However, preharvest food safety is complex as contamination may occur from a variety of environmental sources; such as soil, wildlife feces, and water. Previous studies (9-11) have demonstrated the ability of pathogens to survive for extended periods of time in soil and water, with the potential to contaminate food products. For instance, Baudart et al. (12) reported that river, marine and fresh water sediments were found to contain *Salmonella*, and that sediments could support the ability of *Salmonella* to persist long-term in the environment. This is compounded by the fact that pathogen prevalence in the environment is also dependent on meteorological events. Some studies (13-15) have shown that precipitation increases the detection of pathogens in the environment. For example, Ivanek et al. (13) observed *Listeria* spp. were more prevalent in soil samples when precipitation occurred within two days prior to sample collection. Furthermore, management practices implemented in fields by growers, also impact the risk of produce contamination by pathogens. In one study (16), drip irrigation was shown to reduce the risk of fecal bacterial contamination on plant leaves. Collectively, these findings indicate that every farm has a unique combination of environmental and meteorological factors based on geographical location; in addition to, a specific set of field practices based on landscape and the commodity being cultivated that influence the risk of preharvest contamination.

Consequently, if these factors can be identified within a farm and determinations made on the likelihood of risk then strategies may be devised to break possible transmission pathways, and thereby limit potential contamination events. There is no evidence to suggest that all farms are at equal risk of preharvest contamination; therefore, a general food safety plan would likely

not provide the same risk reduction to all farms. To transition from reactive to preventive risk management, foodborne pathogen contamination in the preharvest environment should be modeled as an ecological process. *Salmonella* and *L. monocytogenes* are known to persist in the environment (6, 10), thereby making them ideal organisms for ecological modeling studies. The following chapters will examine the ecology of *Salmonella* and *L. monocytogenes* in the produce preharvest environment. Specifically, the role of environmental factors and field management practices on pathogen prevalence, persistence, and diversity; in addition, to the potential application of subtyping-based source tracking of pathogens in the preharvest environment.

REFERENCES

1. **Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM.** 2011. Foodborne Illness Acquired in the United States-Major Pathogens. Emerg. Infect. Dis. **17:**7-15.
2. **Painter JA, Hoekstra RM, Ayers T, Tauxe RV, Braden CR, Angulo FJ, Griffin PM.** 2013. Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998-2008. Emerg. Infect. Dis. **19:**407-415.
3. **Mody RK, Greene SA, Gaul L, Sever A, Pichette S, Zambrana I, Dang T, Gass A, Wood R, Herman K, Cantwell LB, Falkenhorst G, Wannemuehler K, Hoekstra RM, McCullum I, Cone A, Franklin L, Austin J, Delea K, Behravesh CB, Sodha SV, Yee JC, Emanuel B, Al-Khaldi SF, Jefferson V, Williams IT, Griffin PM, Swerdlow DL.** 2011. National Outbreak of *Salmonella* Serotype Saintpaul Infections: Importance of Texas Restaurant Investigations in Implicating Jalapeno Peppers. Plos One **6:**7.
4. **Food and Drug Administration.** 2013. Environmental Assessment: Factors Potentially Contributing to the Contamination of Fresh Whole Cantaloupe Implicated in a Multi-State Outbreak of Salmonellosis FDA. http://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm341476.htm.
5. **McCollum JT, Cronquist AB, Silk BJ, Jackson KA, O'Connor KA, Cosgrove S, Gossack JP, Parachini SS, Jain NS, Ettestad P, Ibraheem M, Cantu V, Joshi M, DuVernoy T, Fogg NW, Gorny JR, Mogen KM, Spires C, Teitell P, Joseph LA, Tarr CL, Imanishi M, Neil KP, Tauxe RV, Mahon BE.** 2013. Multistate Outbreak of Listeriosis Associated with Cantaloupe. New Eng. J.Med. **369:**944-953.

6. **Sauders BD, Durak MZ, Fortes E, Windham K, Schukken Y, Lembo AJ, Akey B, Nightingale KK, Wiedmann M.** 2006. Molecular characterization of *Listeria monocytogenes* from natural and urban environments. J. Food Prot. **69:**93-105.

7. **Harris LJ, Farber JN, Beuchat LR, Parish ME, Suslow TV, Garrett EH, Busta FF.** 2003. Outbreaks Associated with Fresh Produce: Incidence, Growth, and Survival of Pathogens in Fresh and Fresh Cut Produce Comprehensive Rev. Food Sci. Food Safety **2:**78-141.

8. **Sant'Ana AS, Barbosa MS, Destro MT, Landgraf M, Franco B.** 2012. Growth potential of *Salmonella* spp. and *Listeria monocytogenes* in nine types of ready-to-eat vegetables stored at variable temperature conditions during shelf-life. Int. J. Food Microbiol. **157:**52-58.

9. **Uesugi AR, Danyluk MD, Mandrell RE, Harris LJ.** 2007. Isolation of *Salmonella* Enteritidis phage type 30 from a single almond orchard over a 5-year period. J. Food Prot. **70:**1784-1789.

10. **Winfield MD, Groisman EA.** 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl. Environ. Microbiol. **69:**3687-3694.

11. **Cooley M, Carychao D, Crawford-Miksza L, Jay MT, Myers C, Rose C, Keys C, Farrar J, Mandrell RE.** 2007. Incidence and Tracking of *Escherichia coli* O157:H7 in a Major Produce Production Region in California. Plos One **2:**16.

12. **Baudart J, Grabulos J, Barusseau JP, Lebaron P.** 2000. *Salmonella* spp. and fecal coliform loads in coastal waters from a point vs. nonpoint source of pollution. J. Environ. Qual. **29:**241-250.

13. **Ivanek R, Grohn YT, Wells MT, Lembo AJ, Sauders BD, Wiedmann M.** 2009. Modeling of Spatially Referenced Environmental and Meteorological Factors Influencing

the Probability of *Listeria* Species Isolation from Natural Environments. Appl. Environ. Microbiol. **75:**5893-5909.

14. **Lipp EK, Schmidt N, Luther ME, Rose JB.** 2001. Determining the effects of El Nino - Southern Oscillation events on coastal water quality. Estuaries **24:**491-497.

15. **Gorski L, Parker CT, Liang A, Cooley MB, Jay-Russell MT, Gordus AG, Atwill ER, Mandrell RE.** 2011. Prevalence, Distribution, and Diversity of *Salmonella enterica* in a Major Produce Region of California. Appl. Environ. Microbiol. **77:**2734-2748.

16. **Bastos RKX, Mara DD.** 1995. The Bacterial Quality of Salad Crops Drip and Furrow Irrigated with Waste Stabilization Pond Effluent. An Evalation of the WHO Guidelines. Water Sci. Technol. **31:**425-430.

CHAPTER 2

LANDSCAPE AND METEOROLOGICAL FACTORS AFFECTING PREVALENCE OF THREE FOODBORNE PATHOGENS IN FRUIT AND VEGETABLE FARMS

Published in: Applied and Environmental Microbiology 79:588-600.

ABSTRACT

Produce related outbreaks have been traced back to the preharvest environment. A longitudinal study was conducted on five farms in New York State to characterize prevalence, persistence, and diversity of foodborne pathogens in fresh produce fields and to determine landscape and meteorological factors that predict their presence. Produce fields were sampled four times per year for two years. A total of 588 samples were analyzed for *L. monocytogenes*, *Salmonella* and Shiga toxin-producing *E. coli* (STEC). The prevalence measures of *L. monocytogenes*, *Salmonella* and STEC were 15.0, 4.6, and 2.7%, respectively. *L. monocytogenes* and *Salmonella* were detected more frequently in water samples, while STEC was detected with equal frequency across all sample types (soil, water, feces and drag swabs). *L. monocytogenes sigB* gene allelic types 57, 58 and 61, and *Salmonella* Cerro, were repeatedly isolated from water samples. Soil available water storage (AWS), temperature, and proximity to three land cover classes: water, roads and urban development, and pasture/hay grass, influenced the likelihood of detecting *L. monocytogenes*. Drainage class, AWS, and precipitation were identified as important factors in *Salmonella* detection. This information was used in a geographic information systems framework to hypothesize locations of environmental reservoirs where the prevalence of foodborne pathogens may be elevated. The map indicated that not all croplands are equally likely to contain environmental reservoirs of *L. monocytogenes*. These findings advance

recommendations to minimize the risk of preharvest contamination by enhancing models of the environmental constraints on the survival and persistence of foodborne pathogens in fields.

INTRODUCTION

Produce safety is an issue of increasing concern. Despite the implementation of produce safety practices, foodborne outbreaks associated with fresh produce continue to result in significant illnesses, hospitalizations and deaths. Approximately 13% of reported foodborne outbreaks were linked to produce from 1990 to 2005 (21). While some of this apparent increase in produce-associated outbreaks may be due to improved surveillance of produce commodities, fresh produce is likely to remain a vehicle for foodborne disease for at least two reasons: i) increased consumption due to promotion of a healthy lifestyle associated with eating fresh produce and ii) fresh produce commodities are often consumed raw with no kill step, such as cooking. As a result, contamination with foodborne pathogens at any point in the supply chain from farm to fork has a heightened chance of reaching the consumer (5). Three major bacterial foodborne pathogens: *Listeria monocytogenes*, Shiga toxin-producing *E. coli* (STEC) and *Salmonella*, have been associated with foodborne disease outbreaks linked to produce. Together, these pathogens account for an estimated 76% (653/861) of deaths attributed to known bacterial foodborne pathogens in the United States (73).

STEC and *Salmonella* have been responsible for the majority of produce-associated foodborne illness outbreaks (21, 36, 77). Many of these outbreaks were traced back to the environmental reservoirs located on the implicated farms. In 2006, one of the first major produce-associated outbreaks linked to preharvest contamination was in spinach. The outbreak-associated strain of *E. coli* O157:H7 was isolated from feral swine, cattle, surface water, sediment and soil from one of the spinach farms (45). An investigation of an outbreak of *Salmonella* Saintpaul in jalapeño peppers recovered the outbreak strain from peppers in the field and from agricultural water used for irrigation from one of the implicated farms in the outbreak (55). While, there are only a few examples of *L. monocytogenes* contamination of produce, *L.*

monocytogenes has the potential to cause produce-associated outbreaks (37, 74). In 2011, a cantaloupe-borne *L. monocytogenes* outbreak caused 146 illnesses, 30 deaths and 1 miscarriage in 28 US states (11). The source of the outbreak was suspected to be a piece of contaminated equipment in the farm packinghouse (27). Although the cause of this outbreak was not due to preharvest contamination of the melons, the results of the outbreak investigation demonstrate the potential risk of *L. monocytogenes* contamination in produce and the difficulties associated with managing this pathogen in the food safety system from farm to fork, because incidental contamination originating from food or environmental sources can persist and amplify through food processing/handling facilities (82).

Foodborne pathogens can enter the environment through many different paths and hosts. Studies have demonstrated the ability of foodborne pathogens to survive for extended periods of time in the soil and water, with the potential to infect new hosts and/or contaminate food products (30, 83, 87). Laboratory and field studies have identified a number of likely sources of preharvest contamination, such as irrigation water, application of untreated manure, runoff water from livestock operations, and wildlife intrusion into fields (7, 8). Management of farms at the farm or production block scale might greatly influence the local movements of the pathogens and the chance for produce to become contaminated. Each farm landscape is a unique combination of numerous environmental characteristics that we hypothesize to set the baseline conditions for persistence of pathogens in or near produce fields. However, contamination of produce in the preharvest environment remains a complex challenge because the conditions that promote persistence of pathogens in the preharvest environment and subsequent produce contamination are not well-understood. The focus of the presented research was to better understand *L. monocytogenes*, STEC and *Salmonella* in the produce preharvest environment and, more specifically, to identify specific, remotely-sensed, topographical properties (e.g., proximity to

forests), soil properties (e.g., available water storage), and meteorological events (e.g., precipitation) that may be associated with their prevalence.

The main objective of the presented research was to assess the prevalence, persistence, and diversity of foodborne pathogens among farms, seasons and sample types (e.g. soil, water, fecal). One hypothesis was that environmental and land management factors control the prevalence of pathogens at spatial scales smaller than a whole farm or season. Classification tree (CT) models were used to identify remotely-sensed landscape (i.e., topographic and edaphic) and meteorological characteristics that delineate the presence and absence of foodborne pathogens in the preharvest environment (18, 44). By modeling foodborne pathogen contamination in the preharvest environment as an ecological process, we seek to supply the produce industry with recommendations to minimize the risk of preharvest contamination.

MATERIALS AND METHODS

Field Sampling Design. A longitudinal field study was performed on five produce farms in New York State. Farms were selected based on the willingness of growers to participate and to sample farms geographically distributed across New York State (NYS). Farms were not selected based on management practices. Farms were sampled nine times from June 2009 to August 2011. Farms were located in three regions of NYS: Central New York (n=1), the Finger Lakes (n=3), and Western New York (n=1). Distance between farms ranged from 33 to 205 km. Sample size calculations were performed using the lower end of the reported range of prevalence estimates for *L. monocytogenes* (84), *Salmonella* (26), and STEC (45) in order to reach 50 isolates for each targeted pathogen. However, due to time and budget constraints, 588 samples were collected that yielded 107, 27 and 16 representative *L. monocytogenes*, *Salmonella*, and STEC isolates,

respectively. Farms were sampled every astronomical season (summer, fall, winter, and spring). Samples were not collected during snow cover in winter.

Within each farm, four fields were selected to standardize sample sizes among farms since the overall farm sizes varied considerably. Fields were selected that had produce commodities generally consumed raw and to capture topographical field diversity such as low and high elevation in the field. During each sampling excursion, a single soil sample was collected, consisting of five subsamples of topsoil from five locations in each field. Soil samples were pooled because pathogens were expected to have high spatial variability and small population sizes across the fields (6, 75). One area drag swab and where available, up to five water and fecal samples were collected for each field. In total, seventy-seven (68 surface and 9 engineered), 9 (all engineered), 45 (44 surface and 1 engineered), 18 (9 surface and 9 engineered), and 25 (all surface) water samples were collected from each of the five farms .Fecal samples represented only 10% of the 588 total samples and the majority of fields did not contain feces. Global positioning system (GPS) coordinates were recorded for each sample collected within the field and revisited upon each subsequent visit, in order to access possible persistence of the targeted foodborne pathogens in the preharvest environment. General farm characteristics were documented (Table 2.1).

Table 2.1 General Farm Characteristics and Key Management

Farm[a]	Size (acres)	Organic	Irrigate[b]	Manure[c]	Compost or composted manure[d]	Staff (no. of employees)	
						Year-round	Temporary
1	> 1000	No	No	Yes	Yes	Yes (31-40)	Yes (40+)
2	< 250	Yes	Yes	Yes	Yes	No	Yes (1-5)
3	> 1000	No	Yes	Yes	No	Yes (6-10)	Yes (21-30)
4	< 250	No	Yes	No	No	No	Yes (11-15)
5	> 1000	No	No	No	No	Yes (11-15)	Yes (31-40)

[a] All farms answered yes to wildlife control measures (e.g., hunting or fences), worker training (e.g., sessions or videos on good hygiene and sanitation practices) and having Good Agricultural Practices (GAP) plans (i.e., third party audits of food safety practices).
[b] Farm 2 used a combination of drip and overhead irrigation depending on crop, farm 3 used overhead irrigation, and Farm 4 uses drip irrigation.
[c] Manure slurry has been applied to a field within the past year. Produce was not planted before 120 days.
[d] Compost and composted manure is a treated product.

Sample Collection. Latex gloves and disposable plastic boot covers (Nasco, Fort Atkinson, WI) were worn for sample collection. Gloves and boot covers were changed between each field and gloves were disinfected with 70% ethanol prior to sample collection. A total of 588 samples were collected. Approximately six inch (15.2 cm) deep soil samples and fecal deposits were gathered into sterile Whirl-Pak bags (Nasco, Fort Atkinson, WI) using sterile scoops (Fisher Scientific, Hampton, NH). Pre-moistened drag swabs, as described by Uesugi *et al.* (83), were dragged around the field perimeter and diagonally back and forth for ≥10 min, covering an average field area of 0.75 ha. Drag swabs were deposited back into the Whirl-Pak sample bag containing 45 mL of phosphate buffered tryptic soy broth (pTSB, Becton Dickinson, Franklin Lakes, NJ). Water samples were collected directly into sterile 250 mL jars by use of a 3.66 m sampling pole (Nasco, Fort Atkinson, WI). These water samples were taken a minimum of 2 m from the shoreline and 0.3 m below the surface. All samples were transported on ice, stored at 4±2°C and processed within 24 h of collection.

Sample Preparation. All samples were used for three separate enrichment schemes to allow for the isolation and detection of (i) *L. monocytogenes*, (ii) *E. coli* O157:H7 and (iii) a combined enrichment for non-O157 STEC and *Salmonella*, as described in detail below. The five soil samples collected in each field were weighed into 5 g portions, and the portions were combined to form a 25 g pooled sample. Three 25 g pooled soil samples were prepared and deposited in sterile filter Whirl-Pak bags for each field. For fecal samples, 2 to 10 g of each fecal sample collected was deposited into three sterile filter Whirl-Pak bags. Drag swab samples were mixed with 90 mL pTSB in the Whirl-Pak bag by hand massaging for 2 min. The drag swab was then squeezed and 10 mL of the liquid contents from the sample bag was aseptically transferred to each of three sterile filter Whirl-Pak bags. Water samples were processed according to Environmental Protection Agency (EPA) standard methods (24, 25). Each water sample

collected (250 mL) was passed through a 0.45 μm filter unit (Nalgene, Rochester, NY). This filter was then aseptically removed and cut into three equal-sized pieces. The pieces of filters were each transferred to a separate sterile filter Whirl-Pak bag.

***L. monocytogenes* enrichment and isolation.** *L. monocytogenes* detection and isolation from environmental samples collected was performed as detailed in previous studies (59, 62, 72). Briefly, samples were diluted 1:10 with buffered *Listeria* enrichment broth (BLEB; Becton Dickinson, Franklin Lakes, NJ). These enrichments were incubated at 30±2°C for 4 h. At 4 h, *Listeria* Selective Enrichment Supplement (Oxoid, Cambridge, UK) was added. After 24 h and 48 h incubations at 30±2°C, 50 μl of each enrichment was streaked onto modified Oxford agar (MOX, Becton Dickinson, Franklin Lakes, NJ) and *L. monocytogenes* plating medium (LMPM, Biosynth International, Itasca, IL). MOX and LMPM plates were incubated for 48 h at 30 and 35±2°C, respectively. Up to ten *L. monocytogenes* presumptive colonies were sub-streaked to brain heart infusion agar (BHI; Becton Dickinson, Franklin Lakes, NJ). BHI agar plates were incubated for 37±2°C for 24 h. Presumptive *L. monocytogenes* colonies were confirmed by polymerase chain reaction (PCR) and partial *sigB* gene sequencing (9, 20, 60).

***E. coli* O157:H7 enrichment and isolation.** Samples in Whirl-pak bags were diluted 1:10 with pTSB. Enrichments were incubated for 2 h at room temperature (23±2°C) to aid in the recovery of injured cells (3). Enrichments were transferred to 42±2°C and incubated for 24 h. Enrichments were subjected to immunomagnetic separation (IMS) to concentrate *E. coli* O157:H7 cells as previously described (56). Washed IMS beads (50 μl) were plated onto two selective and differential media: modified sorbitol-MacConkey agar (mSMAC; Becton Dickinson, Franklin Lakes, NJ) supplemented with 20 mg/L of novobiocin and 2.5 mg/L of potassium tellurite (Sigma-Aldrich, St. Louis, MO) and CHROMagar O157 agar (CHROMagar, Paris, France). CHROMagar O157 and mSMAC plates were incubated at 37±2°C for 24 and 48 h, respectively.

Up to ten presumptive *E. coli* O157:H7 colonies were sub-streaked onto BHI and incubated at 37±2°C for 24 h. Presumptive *E. coli* O157:H7 colonies were confirmed using a multiplex PCR assay that simultaneously screens for *hlyE*, *fliC*$_{h7}$, *eaeA*, *rfbE*, *stx*-I and *stx*-II as previously described (10, 42).

Non-O157 STEC and *Salmonella* enrichment and isolation. The non-selective enrichment step (pTSB) for both pathogens was the same, therefore sample enrichments were sub-sampled. Environmental samples were diluted 1:10 with pTSB. This enrichment was incubated for 2 h at room temperature (23±2 °C), followed by a 24 h incubation at 35±2°C.

To isolate non-O157 STEC, a 1 mL aliquot of the non-selective enrichment was transferred to 9 mL of *E. coli* broth (EC broth; Oxoid) and incubated at 37°C with shaking for 24 h. A 50 µl aliquot of EC broth was plated onto washed sheep's blood agar (Hemostat, Dixon, CA) with 10 mM $CaCl_2$ and 0.5 ug/mL mitomycin C (WBMA; Fisher Scientific, Hampton, NH) and incubated at 35±2°C for 24 h. Up to 20 colonies that demonstrated enterohemolysis were sub-streaked to SMAC plates and incubated at 37±2°C for 24 h. Up to ten colonies that rapidly fermented sorbitol were sub-streaked to BHI and incubated at 37±2°C for 24 h. Presumptive non-O157 STEC colonies were confirmed by the multiplex PCR described above (42) and considered positive if one or both *stx* gene(s) were detected.

Salmonella detection and isolation was performed using a modified version of the procedures outlined in the Food and Drug Administration's Bacteriological Analytical Manual (2). A 1.0 and 0.1 mL aliquot of non-selective pTSB enrichment was transferred to 9 and 9.9 mL of tetrathionate (TT; Oxoid) and Rappaport Vassiliadis (RV; Oxoid; Fisher; Acros Organic, Belgium), respectively. These selective enrichment cultures were incubated in a shaking water bath at 42±2°C for 24 h. A 50 µl aliquot of each selective enrichment was plated onto xylose lysine deoxycholate agar (XLD; Neogen, Lansing, MI) and CHROMagar *Salmonella*

(CHROMagar) agar, and incubated at 35 and 37±2°C for 24 and 48 h, respectively. Up to 20 presumptive *Salmonella* colonies were sub-streaked to BHI and incubated at 37±2°C for 24 h. Presumptive *Salmonella* colonies were confirmed using a previously described PCR assay that detects *invA*, a gene specific to *Salmonella enterica* (49).

Controls and storage. Positive and negative controls were processed in parallel with each pathogen detection and isolation scheme. The following strains were used as positive controls: FSL R3-001 for *Listeria monocytogenes* (*actA* deletion mutant; (70)), ATCC 43895 strain tagged with GFP (FSL F6-825) for *E. coli* O157:H7 (61), FSL F6-704 for non-O157 STEC (*E. coli* O26:H11), and ATCC 700408 strain tagged with GFP (FSL F6-826) for *Salmonella* (56) Negative controls were sterile enrichment media. All isolates were preserved at -80°C in 15% glycerol.

Characterization of Isolates. All *L. monocytogenes*, STEC (*E. coli* O157:H7 and non-O157 STEC), and *Salmonella* isolates were streaked from frozen culture onto BHI and incubated at 37°C for 18 h and a well-isolated colony was selected. Nucleotide sequences of *sigB* from *L. monocytogenes* isolates were obtained by Sanger sequencing performed by the Cornell University Life Sciences Core Laboratories Center and compared with those in the GenBank database using BLASTN to assign allelic types, as defined by a unique combination of polymorphisms (1, 57). *L. monocytogenes* isolates that shared the same allelic type from the same location at least three times were considered possible persistent subtypes and were further subtyped by pulsed field gel electrophoresis (PFGE). PFGE typing was performed using the standard CDC PulseNet protocol using the restriction enzymes *Asc*I and *Apa*I (32). *Salmonella* ser. Braenderup digested with *Xba*I was used as the reference standard, which allowed for normalization and comparison of gel images (43). Pattern images were captured with a Bio-Rad Gel Doc and the Multi Analyst software (Bio-Rad Laboratories, Hercules, CA). PFGE banding

patterns were analyzed using BioNumerics (Applied Maths, Saint-Matins-Latem, Belgium). Comparisons were performed using similarity analyses by using the unweighted pair group-matching algorithm (UPGMA) and the Dice correlation coefficient with a maximum space tolerance of 1.5%.

To confirm the identity of O157:H7 isolates and to determine the serotype of non-O157:H7 STEC isolates, comprehensive O serotyping and H typing were performed on one representative STEC isolate per positive sample at the *E. coli* Reference Center at Pennsylvania State University (State College, PA), as previously described by Orskov et al. (1977) and Machado et al. (2000), respectively (53, 65).

Salmonella cultivation methods use four combinations of selective enrichments and plating media. To account for the possibility of different strains of *Salmonella* being isolated from the sample, one representative isolate per *Salmonella* positive sample from each isolation scheme (e.g., TT to XLD or RV to XLD) was selected for molecular subtyping (54). Serotyping and PFGE were performed on isolates selected. Serotyping, using the White-Kauffman-Le Minor (formerly known as the Kauffman-White scheme) scheme, was performed by the Wadsworth Center, New York State Department of Health (Albany, NY) (33). PFGE typing was performed according to the standard CDC PulseNet protocol for *Salmonella* using the restriction enzyme *Xba*I (69).

Descriptive Data Analysis. Associations of pathogen positive cultures with farm, season, or sample type were determined using a chi-square test. A Fisher's exact test was used if the expected frequency in any cell was less than 5. Confidence intervals were calculated assuming a binominal distribution. Individual *p* values were calculated and were considered statistically significant if less than 0.05. Bonferroni's correction was used to account for multiple testing of the three statistical hypotheses (farm, season and sample type) (41). The diversity of subtypes

within farm, season and sample type was quantified using Simpson's Index of Diversity (D) (76). All statistics for descriptive analyses were performed in SAS 9.1 (SAS Institute Inc., Cary, NC).

Topographical and Soil (spatial) data. Spatially dependent predictor data (e.g., proximity to water and available water storage) were obtained for each sample site (Table S1). Global positioning system (GPS) coordinates of samples were imported into the Geographical Resources Analysis Support System (GRASS) Geographic Information Systems (GIS) environment (79). Site coordinates were re-projected from latitude-longitude into the Universal Transverse Mercator (UTM) coordinate system, North American Datum of 1983. Map layers for land cover (NLCD; National Land Cover Database 2006) and digital elevation model (DEM; Shuttle Radar Topography Mission, 1 arc-second dataset) were acquired from the US Geological Survey (USGS) EarthExplorer geographical databank (http://earthexplorer.usgs.gov/). Map layers for soil characteristics were acquired from the United State Department of Agriculture (USDA) Soil Survey Geographic database (SSURGO) (http://soils.usda.gov/survey/geography/ssurgo/). Road and hydrologic line graphs were obtained from the Cornell University Geospatial Information Repository (CUGIR; http://cugir.mannlib.cornell.edu/). Proximity data were derived from the NLCD land cover basemap by calculating euclidean nearest-neighbor distance to the desired land cover type. Proximity to urban areas was calculated from a map combining road lines with all four classes of developed land cover. Proximity to water was calculated from a map combining water body areas and flow lines. Percent slope was derived from the DEM. In total, 15 different landscape factors were obtained for CT model development, such as soil type; slope; drainage class; available water storage; organic matter; and proximity to urban development, pastures, forests, and water.

Meteorological (temporal) data. For each sample collection date, meteorological variables were obtained from the nearest major airport to each farm, using the airport weather stations in the National Oceanic and Atmospheric Administration (NOAA) National Climate Data Center (NCDC) Local Climatology Database (http://gis.ncdc.noaa.gov/map/lcd/). A major airport was within 60 miles of each farm used in the study. While, small scale differences in weather may be observed between airport and farm, the exploratory nature of the study aimed to capture the association between remotely-sensed meteorological data and pathogen prevalence. In total, 70 different meteorological factors were obtained for CT model development, including temperature (maximum, minimum and daily average) and precipitation amounts. Direct measures of temperature and precipitation were acquired for the day of sampling and three days antecedent. The average temperature and precipitation was calculated for each time period ranging from 1 to 10 days prior to sample collection. Frost cycles were counted by summing the number of times the surface air temperature fluctuated above and below 0°C for each time period ranging from 1 to 10 days prior to sampling. Averaging and frost calculations were performed using a Perl script (Code available from PWB).

Spatial and Temporal Data Analysis. Methods used in our analysis of spatial and temporal factors were adapted from Ivanek et al. (44). Large numbers of landscape and meteorological variables were included in our classification analysis as possible predictors of pathogen presence. Since there was high potential for covariation among landscape and meteorological predictors of pathogen presence, detrending and principal components analysis (PCA) techniques were applied to account for the linear covariation among predictors. PCA was performed using the ade4 package in R 2.13.1 (80).

It was desirable to account for season, temperature and precipitation as independent factors predicting pathogen presence, but season and the meteorological variables did not behave

independently at monthly time scales. Temperature and precipitation were de-trended for the seasonal effect by performing linear regressions and retaining the residuals from these regressions to represent variation of temperature and precipitation within seasons. Soil properties and elevation were also dependent on the general farm properties, so to examine the effects of soil property and elevation variation within farms, these characteristics were de-trended against farm using linear regressions.

De-trended residuals were standardized and used as input for two PCAs to synthesize variation among meteorological and landscape data, respectively, into eigenvectors representing the characteristic behavior of these variables. PCA on meteorological variables yielded an eigenvector that represented 56.1% of the total variation and corresponded well to all temperature variables except average temperature 3 days prior to sampling. The same PCA yielded a second eigenvector describing 18.4% of the total variation that corresponded well to all precipitation variables except precipitation on the day before and the day of sampling. A second PCA showed that landscape data were less likely to be heavily loaded on eigenvectors and demonstrated less covariation among landscape data. This PCA yielded a single useful eigenvector, representing 51% of the total variation. Available water storage and soil organic matter properties were loaded on this eigenvector, but topographic data were retained as independent predictors of pathogen presence. These three eigenvectors were used as predictor variables in CT models, as they synthesized characteristics of multiple, co-varying temperature, precipitation, and soil variables, respectively. This allowed us to minimize the number of predictor variables in the CT models to those that behaved independently.

Classification Tree (CT) Model Development. Tree-based modeling was used to determine rules (i.e., factors) that would classify sampled sites by pathogen presence or absence. Splits were formed by maximizing homogeneity of presence vs. absence results in each node according

to the Gini index (18). The CTs were built using the rpart package in R 2.13.1 (81). To assess predictive power of resulting trees, a cross-validation procedure was performed 25 times for each tree. The detection methods for foodborne pathogens in the environment are not 100% sensitive or specific; therefore the response variable was weighted to maximize the predictive power of the resulting tree. To limit the potential effect of different CT outcomes based on weighting the response variable, we performed a sensitivity analysis in which different weights were applied to negative samples to reflect probabilities of false negatives. The weight of positive samples was always set to 1. The weights for negative results were varied until the weight that minimized cross-validation error was discovered. This weighting scheme was used to produce CTs with 25-fold cross-validation. CTs were pruned to the number of splits that minimized cross-validation error within the selected weighting scheme. This combination of procedures resulted in CTs of reproducible size, predictive power, and split rules and the subsequent analysis of *L. monocytogenes* results by random forest using the CT model supported the CT outcome entirely (data not shown; (16)).

Geospatial search for *L. monocytogenes* reservoirs. Classification trees, and related techniques, result in rules that can be used to predict the most likely areas to observe a species (18), in this case foodborne pathogens. Using *L. monocytogenes* as an example, the pasture proximity and soil property rules from the *L. monocytogenes* CT were applied in a GIS framework to explore the potential for croplands to harbor persistent *L. monocytogenes* in a central New York State landscape. All calculations on maps were performed using GRASS GIS 6.4.1. Raster maps of a) water features and flow lines and b) pasture areas were extended to reflect proximity-based split rules from the CT using the spatial buffering function r.buffer. These rasters were then converted to vector maps, and used in vector map queries using v.overlay to determine cropland areas corresponding to three categories of hypothetical *L.*

monocytogenes reservoir based on CT results. Reservoir polygon areas and minimum reservoir distance from pasture class land areas were calculated using statistical functions in GRASS GIS.

Accession Numbers. Isolate information and subtyping data from this study are archived and available through the Food Microbe Tracker database (http://www.foodmicrobetracker.com) using a guest user login.

RESULTS

***L. monocytogenes* prevalence.** *L. monocytogenes* prevalence was estimated to be 15.0% (88/588) across all samples collected. Farm, season, and sample type were found to be significantly associated with the frequency of *L. monocytogenes* positive samples (Table 2.2). Over the nine collection periods, winter had a consistently higher prevalence of *L. monocytogenes* than all other seasons, the only exception being the summer of 2010. Farm 1 showed a significantly higher prevalence of *L. monocytogenes* compared to Farm 2 (Table 2.2). The prevalence of *L. monocytogenes* was highest amongst water samples (48/174). All *L. monocytogenes* positive water samples were from surface water (e.g., creek or pond); none of the 28 samples from engineered water sources (e.g., municipal or well) were positive for *L. monocytogenes* (Table 2.2).

***Salmonella* prevalence.** The prevalence of *Salmonella* across all samples was 4.6% (27/588). Farm and sample type were significantly associated with the frequency of *Salmonella* positive samples (Table 2.2). While there was no significant seasonal association, *Salmonella* prevalence was greatest in the 2010 and 2011 summers (7.8 and 8.3%, respectively). Farm 1 showed a significantly higher prevalence of *Salmonella* compared to Farms 2 and 5 (Table 2.2) possibly due to the co-management of the produce operations on farm 1 with livestock operations located nearby. The prevalence of *Salmonella* was significantly higher in water samples (16/174)

compared to soil and drag swab samples (4/178 and 3/175, respectively), but similar to fecal samples (4/61) (Table 2.2). All of the 16 *Salmonella* positive water samples originated from surface water (Table 2.2).

STEC prevalence. The prevalence of STEC was 2.7% (16/588) across all samples. Four samples tested positive for *E. coli* O157:H7, including a (i) drag swab sample from a pepper field, (ii) drag swab sample from a sweet corn field, (iii) water sample from a drainage ditch, and (iv) water sample from a creek. None of the factors (e.g., farm, season and sample type) were shown to have a significant association with the frequency of STEC positive samples (Table 2.2). Similar to findings for *L. monocytogenes* and *Salmonella*, all four STEC positive water samples were from surface water (Table 2.2).

Table 2.2 Effect of Factors (Farm, Season and Sample Type) on frequency of positive *L. monocytogenes*, *Salmonella*, and STEC samples found in produce preharvest environments.

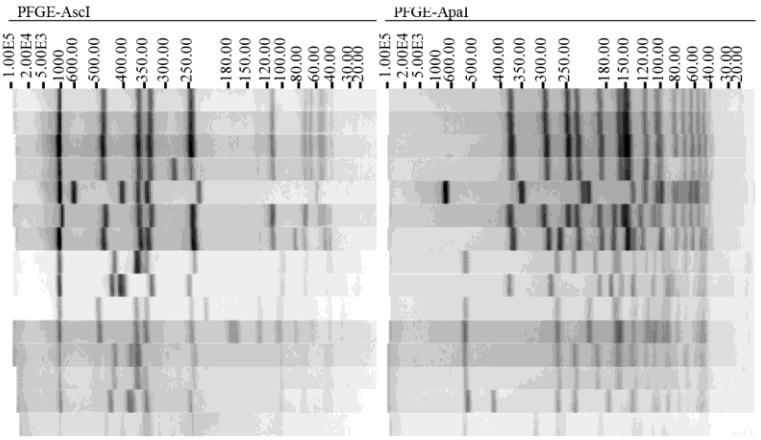
Factor (No. of samples)		Frequency (Percent)[a]		
		L. monocytogenes	*Salmonella*	STEC
Farm				
	1 (166)	39 (23)^{A}	16 (10)^{A}	1 (1)
	2 (103)	5 (5)^{B}	1 (1)^{B}	3 (3)
	3 (113)	13 (12)^{AB}	3 (3)^{AB}	7 (6)
	4 (100)	14 (14)^{AB}	5 (5)^{AB}	1 (1)
	5 (106)	17 (16)^{AB}	2 (2)^{B}	4 (4)
Season				
	Fall (136)	9 (7)^{B}	3 (2)	4 (3)
	Winter (125)	30 (24)^{A}	6 (5)	2 (2)
	Spring (134)	23 (17)^{AB}	6 (4)	3 (2)
	Summer (193)	26 (19)^{A}	12 (9)	7 (5)
Sample Type				
	Soil (178)	16 (9)^{B}	4 (2)^{B}	3 (2)
	Drag Swab (175)	15 (9)^{B}	3 (2)^{B}	5 (3)
	Fecal (61)	9 (15)^{AB}	4 (7)^{AB}	4 (7)
	Water (174)	48 (28)^{A}	16 (9)^{A}	4 (2)
	Engineered (28)	0 (0)	0 (0)	0 (0)
	Surface (146)	48 (33)	16 (11)	4 (3)

[a] Superscript letters represent different statistical populations of values that are significantly different with *P*-value < 0.016. No letters represent values that are not significantly different

.

***L. monocytogenes* diversity.** A total of 107 *L. monocytogenes* isolates were obtained from the collection of 88 environmental samples in which this pathogen was detected. Alignment of *sigB* nucleotide sequences for the 107 isolates showed 12 different *sigB* allelic types. Allelic types belonged to *L. monocytogenes* lineage I, II and IIIa (6, 5, and 1 allelic types, respectively). There was a high diversity of *L. monocytogenes* allelic types amongst farms, seasons, and sample types (D=0.80, 0.78, and 0.85, respectively).

There were four cases of repeat isolation which was defined as the same *sigB* allelic type being isolated three or more times from the same sample site over time. These isolates were subtyped further using PFGE with restriction enzymes *Asc*I and *Apa*I (Fig. 2.1). Analysis of *L. monocytogenes* PFGE showed multiple PFGE patterns for three of the four cases of repeat isolation; however, one case showed identical PFGE patterns for 3 of the 4 isolates obtained from the same water sample site

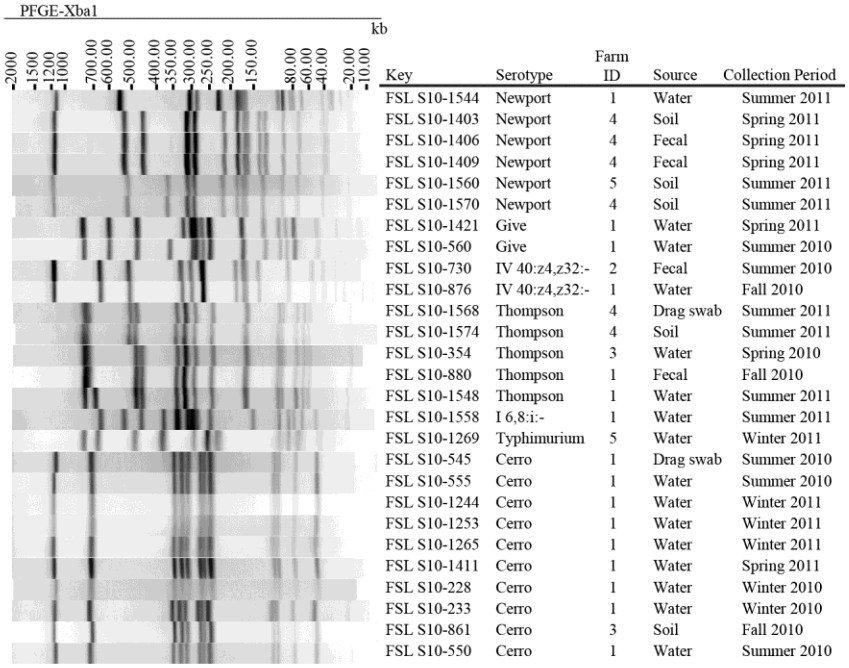


Key	Lineage	Allelic type	Farm ID	Source	Collection Period
FSL S10-009	I	58	1	Water	Summer 2009
FSL S10-084	I	58	1	Water	Fall 2009
FSL S10-366	I	58	1	Water	Spring 2010
FSL S10-598	I	58	1	Water	Summer 2010
FSL S10-020	I	61	1	Water	Summer 2009
FSL S10-086	I	61	1	Water	Fall 2009
FSL S10-601	I	61	1	Water	Summer 2010
FSL S10-298	II	57	1	Water	Winter 2010
FSL S10-604	II	57	1	Water	Summer 2010
FSL S10-1591	II	57	1	Water	Summer 2011
FSL S10-306	II	57	1	Water	Winter 2010
FSL S10-335	II	57	1	Water	Spring 2010
FSL S10-609	II	57	1	Water	Summer 2010
FSL S10-1363	II	57	1	Water	Winter 2011
FSL S10-1490	II	57	1	Water	Spring 2011

Figure 2.1. *Asc*I and *Apa*I PFGE patterns of the four repeat isolation cases of *L. monocytogenes*. The four cases of repeat isolation are as follows: FSL-S10-009, 084, 366 and 598; FSL-S10-020, 086, and 601; FSL-S10-298, 604, and 1591; and FSL-S10-306, 335, 609, 1363, and 1490. In one case (top three PFGE patterns), three of four isolates (FSL-S10-009, 084, and 366) have identical PFGE patterns. Band sizes (kb) are displayed at the top of the PFGE pattern images. PFGE pattern order displayed is result of sample site (manually ordered within BioNumerics).

***Salmonella* diversity**. Serotyping and PFGE was conducted on one representative isolate from each isolation scheme in 26 out of 27 positive samples for a total of 57 *Salmonella* isolates. No isolate was available for typing from one *Salmonella* sample, because preservation failed. One of the 26 available samples yielded a different PFGE type under the four isolation schemes. The two PFGE types were confirmed as *Salmonella* serovars Newport and Thompson. All other isolates from the isolation schemes had identical PFGE types within a sample. The remaining 25 *Salmonella* positive samples contained *Salmonella* serotypes Cerro (10 samples), Newport (5 samples), Thompson (4 samples), Give (2 samples), IV 40:z4,z32:- (2 samples), Typhimurium (1 sample), and I 6,8:i:- (1 sample) (Fig. 2.2). The 7 *Salmonella* serotypes corresponded to eleven different PFGE types (Fig. 2.2). Overall, there was a high level of diversity amongst *Salmonella* serotypes and PFGE types in the produce preharvest environment (serotype D=0.84 and PFGE type D=0.80). *Salmonella* Newport was isolated from two fecal samples and one soil sample from the same field on farm 4 (Fig. 2.2).

Repeat isolation of a *Salmonella* serotype was also observed. *Salmonella* Cerro was isolated from a water source three times during the nine collection periods (Fig. 2.2). The water sample was collected from a creek that was across the road from a field on Farm 1. *Salmonella* Cerro is highly clonal and this particular PFGE pattern matches 89% of *Salmonella* Cerro PFGE patterns (15, 38).



Key	Serotype	Farm ID	Source	Collection Period
FSL S10-1544	Newport	1	Water	Summer 2011
FSL S10-1403	Newport	4	Soil	Spring 2011
FSL S10-1406	Newport	4	Fecal	Spring 2011
FSL S10-1409	Newport	4	Fecal	Spring 2011
FSL S10-1560	Newport	5	Soil	Summer 2011
FSL S10-1570	Newport	4	Soil	Summer 2011
FSL S10-1421	Give	1	Water	Spring 2011
FSL S10-560	Give	1	Water	Summer 2010
FSL S10-730	IV 40:z4,z32:-	2	Fecal	Summer 2010
FSL S10-876	IV 40:z4,z32:-	1	Water	Fall 2010
FSL S10-1568	Thompson	4	Drag swab	Summer 2011
FSL S10-1574	Thompson	4	Soil	Summer 2011
FSL S10-354	Thompson	3	Water	Spring 2010
FSL S10-880	Thompson	1	Fecal	Fall 2010
FSL S10-1548	Thompson	1	Water	Summer 2011
FSL S10-1558	I 6,8:i:-	1	Water	Summer 2011
FSL S10-1269	Typhimurium	5	Water	Winter 2011
FSL S10-545	Cerro	1	Drag swab	Summer 2010
FSL S10-555	Cerro	1	Water	Summer 2010
FSL S10-1244	Cerro	1	Water	Winter 2011
FSL S10-1253	Cerro	1	Water	Winter 2011
FSL S10-1265	Cerro	1	Water	Winter 2011
FSL S10-1411	Cerro	1	Water	Spring 2011
FSL S10-228	Cerro	1	Water	Winter 2010
FSL S10-233	Cerro	1	Water	Winter 2010
FSL S10-861	Cerro	3	Soil	Fall 2010
FSL S10-550	Cerro	1	Water	Summer 2010

Figure 2.2. *Xba*I PFGE patterns of the representative 27 *Salmonella* isolates representing the 26 *Salmonella* positive samples available for typing. One isolate per isolation scheme was PFGE typed; only one representative PFGE pattern is shown. One sample yielded two *Salmonella* PFGE patterns from the four isolation schemes, which represented *Salmonella* Newport and Thompson (FSL-S10-1570 and 1574). Band sizes (kb) are displayed at the top of the PFGE pattern images. PFGE pattern order displayed is result of BioNumerics similarity analyses using the unweighted pair group-matching algorithm (UPGMA) and the Dice correlation coefficient with a maximum space tolerance of 1.5%

.

STEC diversity. Serotyping was conducted to further characterize the 16 STEC isolates. Six distinct O and H serotype results were observed. Serotypes O157:H7 (4/16 isolates) and O8:H19 (4/16 isolates) represented half of the 16 STEC isolates. Additional, serotypes identified were O26:H11 (1/16 isolates), O-:H- (2/16 isolates), OX25:H11 (3/16 isolates), and O91:H49 (2/16 isolates).

Classification of High and Low Prevalence Samples. Classification tree (CT) models were fit using sample presence/absence data in order to further explore the environmental and topographic variables that were associated with the detection of *L. monocytogenes* and *Salmonella* at smaller scales of variation than farm, season or sample type. CTs start with a root node containing all samples and recursively split sample sites by minimizing the mixture between positive and negative environmental samples for the selected foodborne pathogen. CTs often determine multiple possible rules useful for splitting samples. Primary splits exhibited the best improvement score for dividing positive and negative samples into separate nodes; the rule with the second best improvement score was considered a competitor against the primary rule, except in cases where this rule was informationally redundant. In these cases, the next best competitor was selected for display. Surrogate rules represent the predictor that best correlates to the primary rule for the split, and they are used by the algorithm to fill in missing data for the primary rule. These surrogates mimic the primary rule and produce a split with a similar division of positive and negative samples in daughter nodes. No CT was developed for STEC because the trees only produced a root node.

The *L. monocytogenes* CT that gave negative samples one quarter the weight of positive samples resulted in the lowest relative cross-validation error at 0.65 (Fig. 2.3). The CT determined that the prevalence of *L. monocytogenes* in samples collected within 37.5 m of mapped waterways was 39% (29/74). All 74 *L. monocytogenes* samples within 37.5 m of

mapped waterways were surface water samples. All *L. monocytogenes* positive terrestrial samples (n=40) were located farther than 37.5 m from mapped waterways. In samples locations ≥ 37.5 m from water, the eigenvector describing temperature variables split samples such that temperatures that were lower than approximately 2°C below average had 21% prevalence, but samples from warmer temperatures had only 7% prevalence. Using the remotely-sensed average temperature over 5 days prior to sampling, the split rule (Eigenvector_1_Temporal < 1.483) corresponded to < 14°C in summer, < 10°C in spring, or < 5°C in winter. The eigenvector for soil properties included available water storage and soil organic matter. Use of predicted available water storage from the SSURGO database as a representative value revealed soils with available water storage in 0-25 cm depth > 4 cm yielded samples with 31% prevalence versus 10% prevalence in less moist soils. Proximity to pastures was also identified as an important factor in the prediction of *L. monocytogenes* prevalence. Moist terrestrial soil locations sampled at cooler temperatures within 62.5 m of a pasture had 50% (n=25/50) versus 7.5% (3/40) *L. monocytogenes* prevalence in similar sample locations further than 62.5 m from pasture-class land areas. Land use classes were highly interspersed in the areas surrounding sampled farms, so the sample locations meeting the criteria for 50% *L. monocytogenes* prevalence occurred on four of five sampled farms, indicating the proximity to pastures rule was not the product of bias due to the fact that farm 1 had higher prevalence and shorter distances to pasture than the four others. A terminal node was also identified for proximity to urban development or roads (i.e., impervious surface coverage). Locations within 9.5 m of an impervious surface had a predicted *L. monocytogenes* prevalence of 20%, compared to 5% for locations farther than 9.5 m (Fig. 2.3).

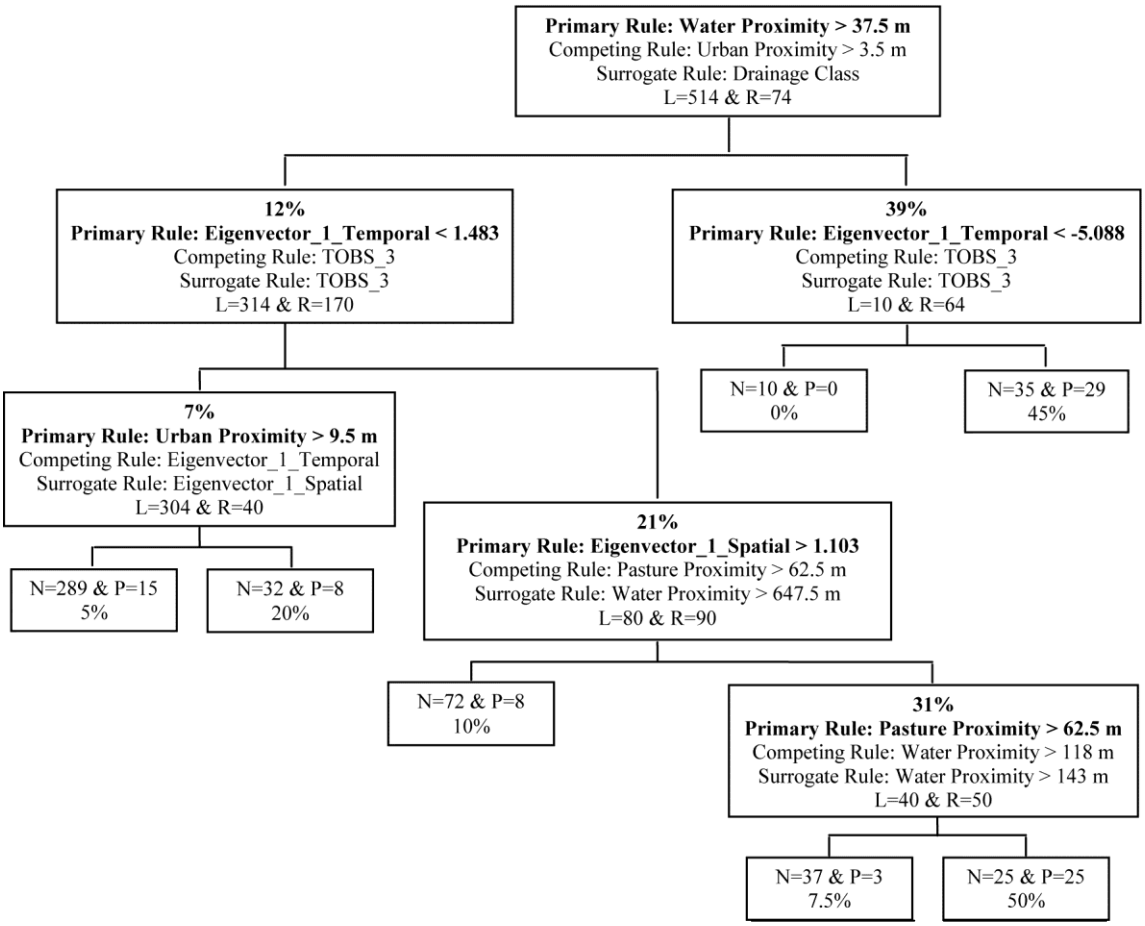


Figure 2.3. CT dividing *L. monocytogenes* environmental samples based on remotely-sensed topographical and meteorological data. On the top of each node there is a rule used for partitioning samples into homogenous subsets. Primary rules are those used to make the depicted split. Percent positive samples is displayed in each node. Rules partition to the left-hand daughter node. Left-hand daughter nodes are enriched for negative samples, and right-hand daughter nodes are enriched for positive samples. Text S3 provides a full summary of the CT. Abbreviations used: L= number of samples partitioned in the left daughter node, R=number of samples partitioned in the right daughter node, N=number of negative samples, and P=number of positive samples.

The *Salmonella* CT that gave negative samples one twentieth the weight of positive samples yielded the lowest relative cross-validation error, 0.67, across all our attempted weighting schemes (Fig 2.4). Drainage class was identified as the most important factor delineating locations of high or low *Salmonella* prevalence. A location where drainage is classified as very poor, somewhat poor, poor, and somewhat drained was determined to have a higher *Salmonella* prevalence (9%) than a location where drainage is classified as moderately well drained and well drained (1.2%). After the tree determined that poorly drained soils contained more *Salmonella*, the algorithm then produced a rule indicating that soils near the upper limit of available water storage were more likely to be negative. *Salmonella* was less likely to occur in soils with available water storage (at 0-25 cm) of 10 cm, which was the maximum value in the soil database. The second temporal eigenvector, which described temporal variation in precipitation, formed another split. In areas with poorly drained soils, *Salmonella* was more prevalent (12%) when measurable precipitation occurred within 3 days prior to sampling.

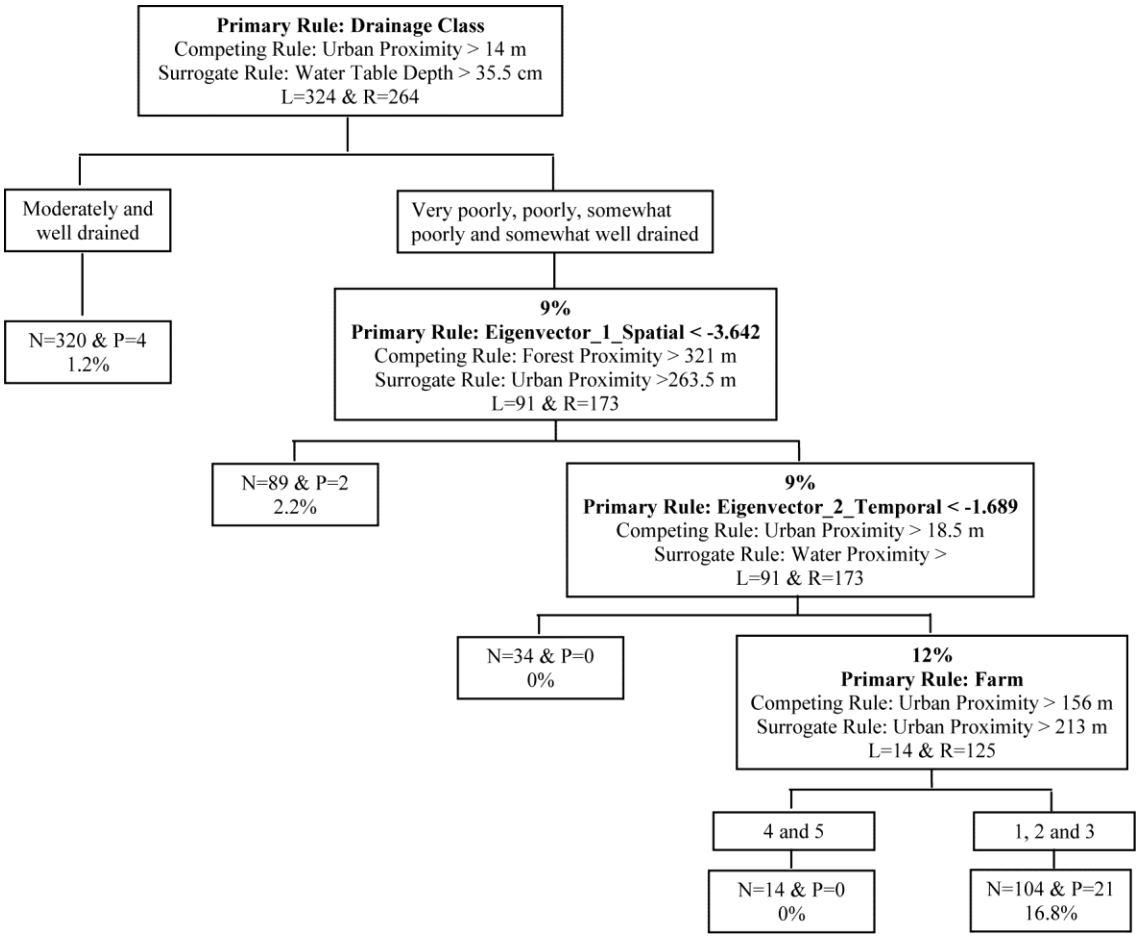


Figure 2.4. CT dividing *Salmonella* environmental samples based on remotely-sensed topographical and meteorological data. On the top of each node there is a rule used for partitioning samples into homogenous subsets. Primary rules are those used to make the depicted split. Percent positive samples is displayed in each node. Rules partition to the left-hand daughter node. Left-hand daughter nodes are enriched for negative samples, and right-hand daughter nodes are enriched for positive samples. Text S4 provides a full summary of the CT. Abbreviations used: L= number of samples partitioned in the left daughter node, R=number of samples partitioned in the right daughter node, N=number of negative samples, and P=number of positive samples.

Geospatial prediction of terrestrial *L. monocytogenes* reservoirs. The CT generated in this study contained topographic and soil property rules that might be useful to map the locations of environmental reservoirs of *L. monocytogenes*. In order to explore the usefulness of the CT rules, a map was generated to represent a CT-based hypothesis about the locations and spatial extents of *L. monocytogenes* environmental reservoirs in an 9024 ha mixed land cover area in central New York State (Fig 2.5). We hypothesize that these locations harbor *L. monocytogenes* within croplands and are more likely to be positive when sampled than other parts of produce fields. Three classes of environmental reservoirs were extracted to produce this map: i) areas within 38 m of mapped surface water, ii) areas outside class (i) with soil available water storage (AWS) > 4.0 cm in the top 25 cm and within 62 m of mapped pasture land cover, and iii) areas with AWS > 4.0 cm in the top 25 cm but outside classes (i) and (ii). All classes were clustered spatially to the west of the main stream in the map, indicating that not all portions of the landscape provide equally good reservoir habitats for *L. monocytogenes*. In the mapped areas, the algorithm identified 221 cropland reservoirs within 38 m of water (light blue; 45% positive samples) and had areas ranging from 0.01 to 4.79 ha with a median of 0.07 ha. One hundred ninety-two reservoirs in higher available water storage soil areas (cyan; 7.5% positive samples), which ranged from 0.1 to 16.2 ha in area with a median of 0.6 ha. Two hundred thirty-five reservoirs within 62 m of pasture class land cover and in soils with higher available water storage (dark blue; 50% positive samples) ranged from 0.1 to 5.2 ha in area with a median area of 0.5 ha.

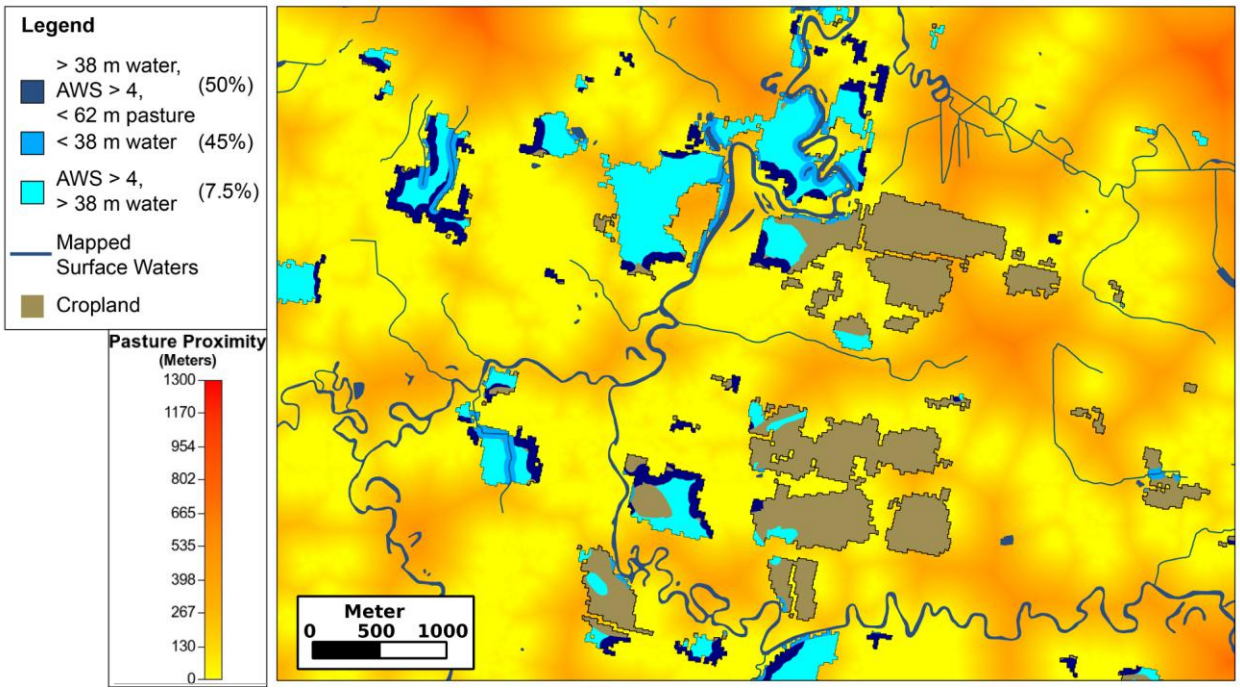


Figure 2.5. Map predicting *L. monocytogenes* environmental reservoir locations and spatial extents based on CT results (Fig. 2.3). Percentages in legend indicate the prevalence of *L. monocytogenes* within that class of reservoir based on the CT. Cropland polygons were isolated were assigned values as reservoirs based on their proximity to water, proximity to pasture and soil available water storage (AWS). The color surface represents proximity to nearest pasture class land cover. Brown cropland polygons are not expected to contain significant *L. monocytogenes* reservoirs.

DISCUSSION

One of the largest practical challenges to ensuring produce safety is to optimize the expenditure of financial and labor resources such that the pathogen will be detected where it is a public health risk. While many aspects of pathogen biology in the environment have been the topic of research (19, 87), field data are lacking on the nature of environmental pathogen reservoirs and how these reservoirs constrain the distribution of pathogens within and among produce fields. As a result, farmers and food safety professionals have little data on which to base sampling schemes that are intended to detect foodborne pathogens in the preharvest environment. This study describes the environmental distribution of three foodborne pathogens at the farm and field scales. We propose that using these analyses provides a means to improve surveillance for foodborne pathogens in produce fields by describing environmental variables that constrain the prevalence of pathogens. These data may also be used to identify areas of high and low predicted pathogen prevalence within farms, enabling more informed decisions about the management of crops associated with foodborne disease outbreaks. While this study does not directly measure the risk of produce contamination, these data can be used to support the development of risk models by describing the environmental microbiology of foodborne pathogens.

The present research was conducted by collecting diverse sample types on privately operated farms. While this has the advantage of enabling the collection of data in farms that are faced with the management challenges and practical considerations of their businesses, it had the disadvantage that sample collection was dictated in large part by convenience to the land owner, so the sampling design quickly became unbalanced and smaller than was originally planned. However, it is important to note that field ecology studies frequently feature unbalanced and under-sampled designs and serious violations of statistical assumptions (18). As an alternative to

violating the assumptions of logistic regression analysis, classification trees are an ideal method to analyze such unbalanced and under-sampled designs (44). The method makes a single, simple, assumption about the distribution of presence/absence observations: that these data can be subdivided according to predictor data in a way that maximizes the homogeneity of the response variable.

Prevalence of key foodborne pathogens is higher in water and fecal samples. Foodborne pathogens are commonly thought to survive better in aquatic and moist extra-host habitats. Water has been identified as both a reservoir and transmission pathway for foodborne pathogen contamination of produce (8, 29). Moreover, all three pathogens examined in this study are known to be common contaminants of agricultural watersheds (35, 51, 52, 64). The range of *L. monocytogenes* and *Salmonella* prevalence in watersheds has been estimated from 6.4 to 62% and 6 to 80%, respectively, based on the region of study (35, 46, 47, 68, 86). A previous study, reported a 7% prevalence of *Salmonella* in water samples obtained from a primary produce growing region located in California (31). Similarly, in the participating NYS produce farms we obtained the highest prevalence of *L. monocytogenes* and *Salmonella* from water samples demonstrating water sources are a potential pathogen reservoir.

However, we recognize there may be a potential bias toward a higher prevalence of pathogens in water samples compared to soil samples in this study. The prevalence of foodborne pathogens in the environment, is expected to be low so there may be a greater chance of detecting a pathogen positive water sample because the pathogen may be distributed more uniformly in water compared to soil samples. The five sub samples of soil collected were also pooled in one composite sample per field, which may contribute to a lower prevalence estimate. Positive *L. monocytogenes* and *Salmonella* samples collected from on-site surface water sources were mostly from small waterways, no more than 3 m wide, as many of the sampled produce

farms are surrounded by developed land (e.g., housing, roadways). Previous research has shown that lower order waterways often receive more direct agricultural drainage and runoff from animal production environments compared to higher order systems, therefore lower order water sources may have elevated pathogen prevalence (51, 52).

The management of water and wildlife on farms may be warranted as evident by our results demonstrating a higher prevalence of *L. monocytogenes* and *Salmonella* in water and feces compared to other environmental sources. Foraging wildlife may also contribute to contamination of fields as demonstrated by isolation of *Salmonella* Newport from soil and wildlife feces (deer) from a single field. Wild and domestic animals are widely known to harbor foodborne pathogens (31, 39, 45). A review of *Salmonella* in wild and domesticated animals (39) determined the prevalence of *Salmonella* in reptiles (6-100%), poultry (0-60%), cattle (2-42%), swine (3.5-28%), rodents (1-15%), birds (3-13%), and domestic cats and dogs (1-5%), which reflects the potential for animals to be sources or vectors of preharvest contamination. However, land and wildlife management changes to prevent wildlife intrusion need to be carefully considered so that new land management practices do not damage the soil, water quality, or endangered/protected species. An optimal solution for the co-management of wildlife habitat and environmental quality is the target of current produce safety research and our research facilitates these efforts by describing the environmental prevalence of pathogens in preharvest environments in relation to potential wildlife habitats on the same landscape (67).

Soil properties and topographic features were identified as constraints on pathogen prevalence in produce fields. The *L. monocytogenes* and *Salmonella* CTs both demonstrate that not all croplands have an equal risk of foodborne pathogen contamination. Soil characteristics and topographic variables corresponding to proximity of sampled areas to other landscape types, including impervious surface (e.g., buildings, roadways), water, or pasture were identified as

factors (i.e., primary rules in the CT) for predicting locations containing pathogens. The primary split for both pathogen CTs were associated with water features, specifically proximity to water and soil drainage class (e.g., well drained, poorly drained) for *L. monocytogenes* and *Salmonella*, respectively. This finding further demonstrates the ability of on-site water sources as potential reservoirs and transmission pathways for foodborne pathogen contamination in the preharvest environment (8, 36, 78).

Proximity to pastures was an important factor influencing the likelihood of detecting a *L. monocytogenes* positive sample. It is important to note that proximity to pastures was obtained through remotely sensed data, and pasture class land-cover can indicate active pasturages, livestock pens, and hay grass fields. It has been shown that livestock shedding and subsequent run-off of foodborne pathogens may be one of the major sources of preharvest contamination (8, 28, 58). Ruminants can shed significant numbers of *L. monocytogenes* while being asymptomatic, and therefore may release large numbers into the environment (15, 59). Additionally, a strong association was shown between the prevalence of *L. monocytogenes* and proximity to cattle and dairy farms in watersheds impacted in agricultural landscapes (51, 72).

Proximity to impervious surfaces was also identified as a factor for the classification of high or low *L. monocytogenes* prevalence locations. Some wildlife carriers of foodborne disease, e.g. rodents and ground nesting birds, use roadside ditches as nesting habitats and may enter croplands from ditches while foraging for food (48). This behavior, particularly in short-dispersal distance species, may cause foodborne pathogen prevalence to be amplified in the edges between residential land and cropland. In addition, impervious surfaces are constructed to remove excess water, usually into bordering drainage ditches, and these may channel fecal contamination from both wildlife and livestock. This contamination could subsequently be spread out of the ditch by local wildlife, heavy precipitation, or human activities.

Soil characteristics, specifically available water storage, soil drainage class, or soil organic matter, were important factors in CTs generated from both *L. monocytogenes* and *Salmonella* prevalence data. Pathogen survival has been shown to increase in moist soils (40, 85). *E. coli* and *Salmonella* held in dry soil for 14 d demonstrated the ability for growth after the soil was moistened with sterile distilled deionized water (12). In addition, it was shown that *E. coli* and *Salmonella* persisted longer in moist soils compared to drier soils. Similarly, *E. coli* numbers were shown to decrease faster in dry soils compared to moist soils (63).

Weather plays a role in pathogen prevalence in the preharvest environment. Meteorological factors were identified in both the *L. monocytogenes* and *Salmonella* CTs, indicating that recent temperature and precipitation can influence foodborne pathogens in the environment, which is consistent with previous findings (31, 35, 44, 71). *L. monocytogenes* was more frequently detected in cooler temperatures that were above freezing prior to sample collection in our study. Two previous studies (34, 86) observed a higher prevalence of *L. monocytogenes* in spring and winter-spring, and proposed it may be due to manure frequently being applied in spring or *L. monocytogenes* being found more frequently on cow hides in the winter-spring, respectively. Ivanek *et al*. (2009) identified freeze/thaw cycles as an important factor in predicting the presence of *Listeria* spp. in vegetation samples (44). Lower frequencies of freeze/thaw cycles before sample collection were predicted to increase the likelihood of *Listeria* spp. detection. *L. monocytogenes* can grow and survive in a wide range of temperatures, but freezing can have an inhibitory effect (22, 23). While temperature was found to be an important factor for predicting the likelihood of detecting a *L. monocytogenes* positive sample in our study and others, it was not found to be a factor in the CT for *Salmonella*. Other studies (35, 71) have identified temperature or season to be associated with frequencies of *Salmonella* in environmental settings; however, it may be concluded the ability to predict an increased likelihood of detecting a positive *Salmonella*

sample, as for other species, may be dependent on local ecology and agriculture practices of the specific location of the study.

Precipitation was identified as a factor influencing the detection of *Salmonella* positive samples in our study. It has been suggested that heavy rain and storm events may aid in the transport of pathogens (50, 66) and potentially lead to higher loads of bacteria in the water along with sediment (4, 14). Foodborne pathogens can survive in sediments for substantial periods, and therefore sediments may be acting as environmental reservoirs harboring pathogens (17). High water flow rates have also been observed to influence pathogen incidence levels in the environment, and may transport pathogens up to 32 km (13). Such long transport distances in waterways may have important implications for the diversity and source tracking of foodborne pathogens impacting preharvest environments (28, 68). It may be necessary to analyze flux of pathogens from potential sources, like livestock pasture, through hydrologic connectivity networks in a dynamic framework that accounts for precipitation in order to accurately estimate risk of *Salmonella* contamination in croplands where it is expected to result from contaminated water sources. The Soil and Water Assessment Tool (SWAT), while unsophisticated in its treatment of bacteria as relatively simple particles, is one tool available to model this process.

Benefits to Predicting the Presence of Foodborne Pathogens in Produce Environments. Transmission of foodborne pathogens to produce in the preharvest environment is a complex process, involving multiple vehicles that transport pathogens from sources (e.g., pasture areas) to sinks (e.g., moist soils in fields), and ultimately to produce. The development of practical tools to predict the presence of pathogens in produce fields or the risk of produce contamination is further complicated by the fact that every produce farm is a unique configuration of spatial and temporal variables. These variables undoubtedly influence the ecology of foodborne pathogens in the environment and may influence the potential for product contamination. Indeed, the

presented analysis indicates that, while the average prevalence of foodborne pathogens in randomly-collected environmental samples is low, local prevalence can be significantly higher under specific combinations of environmental and local land-use conditions. An essential component to developing a mechanistic approach to understanding foodborne pathogen transmission to produce is to characterize environmental reservoirs as favorable or unfavorable pathogen persistence sites. Since CT analysis generates concise rules to delineate pathogen positive and negative sites, application of the resulting rules to remotely-sensed data about farm landscapes can enable the development of specific predictions about expected pathogen presence for any individual produce field. While the CTs presented here require further validation to determine their ultimate usefulness in produce farms, these models advance GIS-enabled modeling to predict the risk of produce contamination. Additionally, fully developed models of how pathogens disperse and persist in the preharvest environment may also permit the development of land management strategies to minimize produce contamination by possibly allowing growers to select crops for these sites that are less susceptible to contamination.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389-3402.
2. **Andrews, W. H., and T. Hammack** 22 March 2012, updated date. Bacteriological Analytical Manual, Chapter 5, *Salmonella*. U.S. Food and Drug Administration, Washington, DC. http://www.fda.gov/Food/ScienceResearch/LaboratoryMethods/BacteriologicalAnalyticalManualBAM/ucm070149.htm.
3. **Barkocy-Gallagher, G. A., E. D. Berry, M. Rivera-Betancourt, T. M. Arthur, X. W. Nou, and M. Koohmaraie.** 2002. Development of methods for the recovery of *Escherichia coli* O157 : H7 and *Salmonella* from beef carcass sponge samples and bovine fecal and hide samples. J. Food Prot. **65:**1527-1534.
4. **Baudart, J., J. Grabulos, J. P. Barusseau, and P. Lebaron.** 2000. *Salmonella* spp. and fecal coliform loads in coastal waters from a point vs. nonpoint source of pollution. J. Environ. Qual. **29:**241-250.
5. **Berger, C. N., S. V. Sodha, R. K. Shaw, P. M. Griffin, D. Pink, P. Hand, and G. Frankel.** 2010. Fresh fruit and vegetables as vehicles for the transmission of human pathogens. Environ. Microbiol. **12:**2385-2397.
6. **Bergholz PW, Noar JD, Buckley DH.** 2011. Environmental patterns are imposed on the population structure of *Escherichia coli* after fecal deposition. Appl. Environ. Microbiol. **77:**211-219.
7. **Beuchat, L. R.** 1996. Pathogenic microorganisms associated with fresh produce. J. Food Prot. **59:**204-216.

8. **Beuchat, L. R.** 2006. Vectors and conditions for preharvest contamination of fruits and vegetables with pathogens capable of causing enteric diseases. Br. Food J. **108:**38-53.

9. **Bundrant, B. N., T. Hutchins, H. C. den Bakker, E. Fortes, and M. Wiedmann.** 2011. Listeriosis outbreak in dairy cattle caused by an unusual *Listeria monocytogenes* serotype 4b strain. J. Vet. Diagn. Invest. **23:**155-158.

10. **Carlson, B. A., K. K. Nightingale, J. A. Scanga, J. D. Tatum, J. N. Sofos, G. C. Smith, and K. E. Belk.** 2006. Identification and evaluation of cattle persistently shedding vs. cattle non-persistently shedding *Escherichia coli* O157:H7. Final Report.

11. **Centers for Disease Control and Prevention.** 08 December 2011, posting date. Final Report: Multistate outbreak of Listeriosis linked to whole cantaloupes from Jensen farms, Colorado. Centers for Disease Control and Prevention. Atlanta, GA. http://www.cdc.gov/listeria/outbreaks/cantaloupes-jensen-farms/120811/index.html

12. **Chandler, D. S., and J. A. Craven.** 1980. Relationship of soil-soil-moisture to survival of *Escherichia coli* and *Salmonella* Typhimurium in soils. Aust. J. Agr. Res. **31:**547-555.

13. **Cooley, M., D. Carychao, L. Crawford-Miksza, M. T. Jay, C. Myers, C. Rose, C. Keys, J. Farrar, and R. E. Mandrell.** 2007. Incidence and tracking of *Escherichia coli* O157:H7 in a major produce production region in California. PLoS One **2:**16.

14. **Crabill, C., R. Donald, J. Snelling, R. Foust, and G. Southam.** 1999. The impact of sediment fecal coliform reservoirs on seasonal water quality in Oak Creek, Arizona. Water Res. **33:**2163-2171.

15. **Cummings, K. J., L. D. Warnick, M. Elton, L. D. Rodriguez-Rivera, J. D. Siler, E. M. Wright, Y. T. Groehn, and M. Wiedmann.** 2010. *Salmonella enterica* serotype Cerro among dairy cattle in New York: an emerging pathogen? Foodborne Pathog. Dis. **7:**659-665.

16. **Cutler, D. R., T. C. Edwards, Jr., K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler.** 2007. Random forests for classification in ecology. Ecology **88:**2783-2792.

17. **Czajkowska, D., A. Witkowska-Gwiazdowska, I. Sikorska, H. Bowczyk-Maleszak, and M. Horoch.** 2005. Survival of *Escherichia coli* serotype O157:H7 in water and in bottom-shore sediments. Pol. J. Environ. Stud. **14:**423-430.

18. **De'ath, G., and K. E. Fabricius.** 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology **81:**3178-3192.

19. **Deering AJ, Mauer LJ, Pruitt RE.** 2012. Internalization of *E. coli* O157:H7 and *Salmonella* spp. in plants: A review. Food Res. Int. **45:**567-575.

20. **den Bakker, H. C., B. N. Bundrant, E. D. Fortes, R. H. Orsi, and M. Wiedmann.** 2010. A population genetics-based and phylogenetic approach to understanding the evolution of virulence in the genus *Listeria*. Appl. Environ. Microbiol. **76:**6085-6100.

21. **Doyle, M. P., and M. C. Erickson.** 2008. Summer meeting 2007 - the problems with fresh produce: an overview. J. Appl. Microbiol. **105:**317-330.

22. **Elkest, S. E., and E. H. Marth.** 1992. Freezing of *Listeria monocytogenes* and other microorganisms - a review. J. Food Prot. **55:**639-648.

23. **Elkest, S. E., A. E. Yousef, and E. H. Marth.** 1991. Fate of *Listeria monocytogenes* during freezing and frozen storage. J. Food Sci. **56:**1068-1071.

24. **Environmental Protection Agency.** 2000. Improved enumeration methods for the recreational water quality indicators: *Enterococci* and *Escherichia coli*, vol. EPA/821/R-97/004. Office or Science and Technology, Washington, D.C.

25. **Environmental Protection Agency.** 2002. Method 1603: *Escherichia coli* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar (modified mTEC), vol. EPA 821-R-02-023. Office of Science and Technology,

Washington D.C.

26. **Espinoza-Medina IE, Rodriguez-Leyva FJ, Vargas-Arispuro I, Islas-Osuna MA, Acedo-Felix E, Martinez-Tellez MA.** 2006. PCR identification of *Salmonella*: Potential contamination sources from production and postharvest handling of cantaloupes. J. Food Prot. **69:**1422-1425.

27. **Food and Drug Administration.** 09 January 2012, posting date. Final Update on Multistate Outbreak of Listeriosis Linked to Whole Cantaloupes. Food and Drug Administration. Silver Spring, MD. http://www.fda.gov/food/foodsafety/corenetwork/ucm272372.htm

28. **Ferguson, C., A. M. D. Husman, N. Altavilla, D. Deere, and N. Ashbolt.** 2003. Fate and transport of surface water pathogens in watersheds. Crit. Rev. Environ. Sci. Technol. **33:**299-361.

29. **Foltz, V. D.** 1969. *Salmonella* ecology. J. Am. Oil Chem. Soc. **46:**222-224.

30. **Fremaux, B., C. Prigent-Combaret, and C. Vernozy-Rozand.** 2008. Long-term survival of shiga toxin-producing *Escherichia coli* in cattle effluents and environment: an updated review. Vet. Microbiol. **132:**1-18.

31. **Gorski, L., C. T. Parker, A. Liang, M. B. Cooley, M. T. Jay-Russell, A. G. Gordus, E. R. Atwill, and R. E. Mandrell.** 2011. Prevalence, distribution, and diversity of *Salmonella enterica* in a major produce region of California. Appl. Environ. Microbiol. **77:**2734-2748.

32. **Graves, L. M., and B. Swaminathan.** 2001. PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. Int. J. Food Microbiol. **65:**55-62.

33. **Grimont, P. A. D., and F. X. Weill.** 2007. Antigenic formulae of the *Salmonella*

Serovars, 9th ed. World Health Organization Collaborating Centre for Reference and Research on *Salmonella*, Pasteur Institute, Paris.

34. **Guerini, M. N., D. M. Brichta-Harhay, S. D. Shackelford, T. M. Arthur, J. M. Bosilevac, N. Kalchayanand, T. L. Wheeler, and M. Koohmaraie.** 2007. *Listeria* prevalence and *Listeria monocytogenes* serovar diversity at cull cow and bull processing plants in the united states. J. Food Prot. **70:**2578-2582.

35. **Haley, B. J., D. J. Cole, and E. K. Lipp.** 2009. Distribution, diversity, and seasonality of waterborne *Salmonellae* in a rural watershed. Appl. Environ. Microbiol. **75:**1248-1255.

36. **Hanning, I. B., J. D. Nutt, and S. C. Ricke.** 2009. Salmonellosis outbreaks in the United States due to fresh produce: sources and potential intervention measures. Foodborne Pathog. Dis. **6:**635-648.

37. **Harris, L. J., J. N. Farber, L. R. Beuchat, M. E. Parish, T. V. Suslow, E. H. Garrett, and F. F. Busta.** 2003. Outbreaks associated with fresh produce: incidence, growth, and survival of pathogens in fresh and fresh cut produce. Compr. Rev. Food Sci. Food Safety. **2:**78-141.

38. **Hoelzer, K., K. J. Cummings, E. M. Wright, L. D. Rodriguez-Rivera, S. E. Roof, A. I. M. Switt, N. Dumas, T. Root, D. J. Schoonmaker-Bopp, Y. T. Grohn, J. D. Slier, L. D. Warnick, D. D. Hancock, M. A. Davis, and M. Wiedmann.** 2011. *Salmonella* Cerro isolated over the past twenty years from various sources in the US represent a single predominant pulsed-field gel electrophoresis type. Vet. Microbiol. **150:**389-393.

39. **Hoelzer, K., A. I. M. Switt, and M. Wiedmann.** 2011. Animal contact as a source of human non-typhoidal Salmonellosis. Vet Res **42**.

40. **Holley, R. A., K. M. Arrus, K. H. Ominski, M. Tenuta, and G. Blank.** 2006. *Salmonella* survival in manure-treated soils during simulated seasonal temperature

exposure. J. Environ. Qual. **35:**1170-1180.

41. **Holm, S., S. Mark, and T. Adolfsson.** 2005. A step-down test for effects in unreplicated factorial designs. Commun. Stat-Theor. M **34:**405-416.

42. **Hu, Y., Q. Zhang, and J. C. Meitzler.** 1999. Rapid and sensitive detection of *Escherichia coli* O157:H7 in bovine faeces by a multiplex PCR. J. Appl. Microbiol. **87:**867-876.

43. **Hunter, S. B., P. Vauterin, M. A. Lambert-Fair, M. S. Van Duyne, K. Kubota, L. Graves, D. Wrigley, T. Barrett, and E. Ribot.** 2005. Establishment of a universal size standard strain for use with the PulseNet standardized pulsed-field gel electrophoresis protocols: converting the national databases to the new size standard. J. of Clin. Microbiol. **43:**1045-1050.

44. **Ivanek, R., Y. T. Grohn, M. T. Wells, A. J. Lembo, B. D. Sauders, and M. Wiedmann.** 2009. Modeling of spatially referenced environmental and meteorological factors influencing the probability of *Listeria* species isolation from natural environments. Appl. Environ. Microbiol. **75:**5893-5909.

45. **Jay, M. T., M. Cooley, D. Carychao, G. W. Wiscomb, R. A. Sweitzer, L. Crawford-Miksza, J. A. Farrar, D. K. Lau, J. O'Connell, A. Millington, R. V. Asmundson, E. R. Atwill, and R. E. Mandrell.** 2007. *Escherichia coli* O157:H7 in feral swine near spinach fields and cattle, central California coast. Emerg. Infect. Dis **13:**1908-1911.

46. **Johnson, J. Y. M., J. E. Thomas, T. A. Graham, I. Townshend, J. Byrne, L. B. Selinger, and V. P. J. Gannon.** 2003. Prevalence of *Escherichia coli* O157:H7 and *Salmonella* spp. in surface waters of southern Alberta and its relation to manure sources. Can. J. Microbiol. **49:**326-335.

47. **Jokinen, C., T. A. Edge, S. Ho, W. Koning, C. Laing, W. Mauro, D. Medeiros, J.**

Miller, W. Robertson, E. Taboada, J. E. Thomas, E. Topp, K. Ziebell, and V. P. J. Gannon. 2011. Molecular subtypes of *Campylobacter* spp., *Salmonella enterica*, and *Escherichia coli* O157:H7 isolated from faecal and surface water samples in the Oldman River watershed, Alberta, Canada. Water Res. **45:**1247-1257.

48. **Kaufman, D. W., G. A. Kaufman, and B. K. Clark.** 2000. Small mammals in native and anthropogenic habitats in the Lake Wilson area of north-central Kansas. Southwest Nat. **45:**45-60.

49. **Kim, J. S., G. G. Lee, J. S. Park, Y. H. Jung, H. S. Kwak, S. B. Kim, Y. S. Nam, and S. T. Kwon.** 2007. A novel multiplex PCR assay for rapid and simultaneous detection of five pathogenic bacteria: *Escherichia coli* O157:H7, *Salmonella*, *Staphylococcus aureus, Listeria monocytogenes*, and *Vibrio parahaemolyticus*. J. Food Prot. **70:**1656-1662.

50. **Lipp, E. K., N. Schmidt, M. E. Luther, and J. B. Rose.** 2001. Determining the effects of El Nino - southern oscillation events on coastal water quality. Estuaries **24:**491-497.

51. **Lyautey, E., D. R. Lapen, G. Wilkes, K. McCleary, F. Pagotto, K. Tyler, A. Hartmann, P. Piveteau, A. Rieu, W. J. Robertson, D. T. Medeiros, T. A. Edge, V. Gannon, and E. Topp.** 2007. Distribution and characteristics of *Listeria monocytogenes* isolates from surface waters of the South Nation River watershed, Ontario, Canada. Appl. Environ. Microbiol. **73:**5401-5410.

52. **Lyautey, E., Z. Lu, D. R. Lapen, G. Wilkes, A. Scott, T. Berkers, T. A. Edge, and E. Topp.** 2010. Distribution and diversity of *Escherichia coli* populations in the South Nation River drainage basin, eastern Ontario, Canada. Appl. Environ. Microbiol. **76:**1486-1496.

53. **Machado, J., F. Grimont, and P. A. D. Grimont.** 2000. Identification of *Escherichia coli* flagellar types by restriction of the amplified *fliC* gene. Res. Microbiol. **151:**535-546.

54. **Michael, G. B., R. Simoneti, M. da Costa, and M. Cardoso.** 2003. Comparison of different selective enrichment steps to isolate *Salmonella* spp. from feces of finishing swine. Braz. J. Microbiol. **34:**138-142.

55. **Mody, R. K., S. A. Greene, L. Gaul, A. Sever, S. Pichette, I. Zambrana, T. Dang, A. Gass, R. Wood, K. Herman, L. B. Cantwell, G. Falkenhorst, K. Wannemuehler, R. M. Hoekstra, I. McCullum, A. Cone, L. Franklin, J. Austin, K. Delea, C. B. Behravesh, S. V. Sodha, J. C. Yee, B. Emanuel, S. F. Al-Khaldi, V. Jefferson, I. T. Williams, P. M. Griffin, and D. L. Swerdlow.** 2011. National outbreak of *Salmonella* serotype Saintpaul infections: importance of Texas restaurant investigations in implicating jalapeno peppers. PLoS One **6:**7.

56. **Mukherjee, A., D. Speh, A. T. Jones, K. M. Buesing, and F. Diez-Gonzalez.** 2006. Longitudinal microbiological survey of fresh produce grown by farmers in the upper midwest. J. Food Prot. **69:**1928-1936.

57. **Nightingale, K., L. Bovell, A. Grajczyk, and M. Wiedmann.** 2007. Combined *sigB* allelic typing and multiplex PCR provide improved discriminatory power and reliability for *Listeria monocytogenes* molecular serotyping. J. Microbiol. Methods **68:**52-59.

58. **Nightingale, K. K., E. D. Fortes, C. R. Nightingale, Z. Her, Y. H. Schukken, Y. T. Grohn, and M. Wiedmann.** 2003. Transmission and ecology of *Listeria monocytogenes* in ruminant hosts and the pre-harvest food environment, abstr P-079. Abstr. 103rd Gen. Meet. Am. Soc. Microbiol. American Society for Microbiology, Washington, DC.

59. **Nightingale, K. K., Y. H. Schukken, C. R. Nightingale, E. D. Fortes, A. J. Ho, Z. Her, Y. T. Grohn, P. L. McDonough, and M. Wiedmann.** 2004. Ecology and transmission of *Listeria monocytogenes* infecting ruminants and in the farm environment. Appl. Environ. Microbiol. **70:**4458-4467.

60. **Nightingale, K. K., K. Windham, and M. Wiedmann.** 2005. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal Listeriosis cases and foods. J. Bacteriol. **187:**5537-5551.

61. **Noah, C. W., C. I. Shaw, J. S. Ikeda, K. S. Kreuzer, and J. N. Sofos.** 2005. Development of green fluorescent protein-expressing bacterial strains and evaluation for potential use as positive controls in sample analyses. J. Food Prot. **68:**680-686.

62. **Norton, D. M., M. A. McCamey, K. L. Gall, J. M. Scarlett, K. J. Boor, and M. Wiedmann.** 2001. Molecular studies on the ecology of *Listeria monocytogenes* in the smoked fish processing industry. Appl. Environ. Microbiol. **67:**198-205.

63. **Ohtomo, R., K. Minato, and M. Saito.** 2004. Survival of *Escherichia coli* in a field amended with cow feces slurry. Soil Sci. Plant Nutr. **50:**575-581.

64. **Oliver, D. M., C. D. Clegg, P. M. Haygarth, and A. L. Heathwaite.** 2005. Assessing the potential for pathogen transfer from grassland soils to surface waters, p. 125-180. *In* D. L. Sparks (ed.), Advances in Agronomy, Vol 85, vol. 85. Elsevier Academic Press Inc, San Diego.

65. **Orskov, I., F. Orskov, B. Jann, and K. Jann.** 1977. Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriol. Rev. **41:**667-710.

66. **Oshea, M. L., and R. Field.** 1992. Detection and Disinfection of pathogens in storm generated flows. Can. J. Microbiol. **38:**267-276.

67. **Parker, J. S., R. S. Wilson, J. T. LeJeune, L. Rivers, and D. Doohan.** 2012. An expert guide to understanding grower decisions related to fresh fruit and vegetable contamination prevention and control. Food Control **26:**107-116.

68. **Patchanee, P., B. Molla, N. White, D. E. Line, and W. A. Gebreyes.** 2010. Tracking *Salmonella* contamination in various watersheds and phenotypic and genotypic diversity.

Foodborne Pathog. Dis. **7:**1113-1120.

69. **Ribot, E. M., M. A. Fair, R. Gautom, D. N. Cameron, S. B. Hunter, B. Swaminathan, and T. J. Barrett.** 2006. Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* for PulseNet. Foodborne Pathog. Dis. **3:**59-67.

70. **Roberts, A. J., and M. Wiedmann.** 2006. Allelic exchange and site-directed mutagenesis probe the contribution of *ActA* amino-acid variability to phosphorylation and virulence-associated phenotypes among *Listeria monocytogenes* strains. Fems. Microbiol. Lett. **254:**300-307.

71. **Rodriguez, A., P. Pangloli, H. A. Richards, J. R. Mount, and F. A. Draughon.** 2006. Prevalence of *Salmonella* in diverse environmental farm samples. J. Food Prot. **69:**2576-2580.

72. **Sauders, B. D., M. Z. Durak, E. Fortes, K. Windham, Y. Schukken, A. J. Lembo, B. Akey, K. K. Nightingale, and M. Wiedmann.** 2006. Molecular characterization of *Listeria monocytogenes* from natural and urban environments. J. Food Prot. **69:**93-105.

73. **Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin.** 2011. Foodborne illness acquired in the United States major pathogens. Emerg. Infect. Dis **17:**7-15.

74. **Schlech, W. F., P. M. Lavigne, R. A. Bortolussi, A. C. Allen, E. V. Haldane, A. J. Wort, A. W. Hightower, S. E. Johnson, S. H. King, E. S. Nicholls, and C. V. Broome.** 1983. Epidemic Listeriosis - evidence for transmission by food. N. Engl. J. Med. **308:**203-206.

75. **Semenov AV, Franz E, van Overbeek L, Termorshuizen AJ, van Bruggen AH.** 2008. Estimating the stability of *Escherichia coli* O157:H7 survival in manure-amended soils

with different management histories. Environ. Microbiol. **10:**1450-1459.

76. **Simpson, E. H.** 1949. Measurement of diversity. Nature **163:**688-688.

77. **Sivapalasingam, S., C. R. Friedman, L. Cohen, and R. V. Tauxe.** 2004. Fresh produce: a growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. J. Food Prot. **67:**2342-2353.

78. **Tauxe, R., H. Kruse, C. Hedberg, M. Potter, J. Madden, and K. Wachsmuth.** 1997. Microbial hazards and emerging issues associated with produce - a preliminary report to the National Advisory Committee on Microbiologic Criteria for Foods. J. Food Prot. **60:**1400-1408.

79. **GRASS Development Team.** 2012 Geographic Resources Analysis Support System (GRASS) Software, Version 6.4.1. Open Source Geospatial Foundation. http://grass.osgeo.org

80. **R Development Core Team.** 2012**.** R: A language and environment for statistical computing. R Foundation for Statistical Computing, http://www.R-project.org/.

81. **Therneau, T.M., B. Atkinson.** 2012. rpart: Recursive Partitioning. R package version 3.1-52. http://mayoresearch.mayo.edu/mayo/research/biostat/splusfunctions.cfm

82. **Thimothe, J., K. K. Nightingale, K. Gall, V. N. Scott, and M. Wiedmann.** 2004. Tracking of *Listeria monocytogenes* in smoked fish processing plants. J. Food Prot. **67:**328-341.

83. **Uesugi, A. R., M. D. Danyluk, R. E. Mandrell, and L. J. Harris.** 2007. Isolation of *Salmonella* Enteritidis phage type 30 from a single almond orchard over a 5-year period. J. Food Prot. **70:**1784-1789.

84. **Weis J, Seeliger HPR.** 1975. Incidence of *Listeria monocytogenes* in nature. Appl. Microbiol. **30:**29-32.

85. **Welshimer, H. J.** 1960. Survival of *Listeria monocytogenes* in soil. J. Bacteriol. **80:**316-320.

86. **Wilkes, G., T. A. Edge, V. P. J. Gannon, C. Jokinen, E. Lyautey, N. F. Neumann, N. Ruecker, A. Scott, M. Sunohara, E. Topp, and D. R. Lapen.** 2011. Associations among pathogenic bacteria, parasites, and environmental and land use factors in multiple mixed-use watersheds. Water Res. **45:**5807-5825.

87. **Winfield, M. D., and E. A. Groisman.** 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl. Environ. Microbiol. **69:**3687-3694.

CHAPTER 3

RISK FACTORS ASSOCIATED WITH *SALMONELLA* AND *LISTERIA MONOCYTOGENES* CONTAMINATION OF PRODUCE FIELDS

Published in: Applied and Environmental Microbiology **79**:7618-7627

ABSTRACT

Identification of management practices associated with preharvest pathogen contamination of produce fields is crucial to the development of effective Good Agricultural Practices (GAPs). A cross-sectional study was conducted to (i) determine management practices associated with a *Salmonella* or *Listeria monocytogenes*-positive field and (ii) quantify the frequency of these pathogens in irrigation and non-irrigation water sources. Over five weeks, 21 produce farms in New York State were visited. Field-level management practices were recorded for 263 fields, and 600 environmental samples (soil, drag swab, and water) were collected and analyzed for *Salmonella* and *L. monocytogenes*. Management practices were evaluated for their association with the presence of a pathogen-positive field. *Salmonella* and *L. monocytogenes* were detected in 6.1% and 17.5% of fields (n=263), and 11% and 30% of water samples (n=74), respectively. The majority of pathogen-positive water samples were from non-irrigation surface water sources. Multivariate analysis showed that manure application within a year increased the odds of a *Salmonella*-positive field (odds ratio [OR] 16.7), while presence of a buffer zone had a protective effect (OR 0.1). Irrigation (within 3 days of sample collection, OR 6.0), reported wildlife observation (within 3 days of sample collection, OR 6.1), and soil cultivation (within 7 days of sample collection, OR 2.9) all increased the likelihood of an *L. monocytogenes*-positive field. Our findings provide new data that will assist growers with science-based evaluation of

their current GAPs and implementation of preventive controls that reduce the risk of preharvest contamination.

INTRODUCTION

Produce commodities have been estimated to account for an estimated 46%, 38% and 23% of foodborne illnesses, hospitalizations and deaths in the United States (US), respectively (1). The fact that produce commodities are often consumed raw or with minimal processing likely contributes to the risk of foodborne disease associated with produce. *Salmonella* and *Listeria monocytogenes* are two bacterial foodborne pathogens that represent a substantial burden to the produce industry. Produce-borne *Salmonella* outbreaks have been responsible for a considerable number of foodborne illness cases (2-6). For example, a *Salmonella* outbreak in 2005, associated with tomatoes, resulted in 459 illnesses across 21 US states (7). In 2008, an outbreak of *Salmonella*, linked to jalapeno peppers, sickened approximately 1500 individuals from 43 states, the "District of Colombia" and Canada; this became the largest known outbreak of foodborne illness in the US within the past decade (8). *L. monocytogenes* was responsible for a 2011 produce-borne outbreak, in the US, with 147 illnesses, 33 deaths and 1 miscarriage, due to consumption of cantaloupe (9). In addition, a considerable number of produce recalls have occurred in the past three years as a result of *L. monocytogenes* contamination (e.g., spinach, lettuce (10)). Both *Salmonella* and *L. monocytogenes* can contaminate, persist, and amplify at any point along the farm-to-fork continuum from production to consumption; therefore, minimizing the risk of contamination by these pathogens throughout the supply chain is essential to reducing foodborne illness risks (11-13).

The risk of produce contamination can be reduced by controlling for conditions that favor pathogen introduction and growth in the preharvest environment. Preharvest produce safety is complicated by the fact that each farm has a distinct combination of environmental risk factors (e.g., topography, land-use interactions, and climate). Combinations of these environmental factors influence the frequency and transmission of foodborne pathogens, and subsequently

impact the risk of produce contamination (14). Mitigating contamination risks from environmental factors may be complex and challenging (e.g., as it is difficult to modify farm landscapes); however, modifying management practices to minimize contamination risks may be a more achievable approach. Eighty-nine percent of growers in the US have already reported implementing at least one on-farm food safety measure due to pressure from auditors, inspectors, buyers, and other food safety professionals (15). Examples of food safety measures that were implemented include removing riparian areas, treating irrigation water, installing fences, and using poison bait to control rodents. While these practices were initially used to limit food safety risks in high risk crops (e.g., leafy greens, tomatoes), a follow-up study determined these practices were also being applied to low risk crops (e.g., potatoes, squash), thus increasing the cost of production (16). In addition, some of these practices may also have negative effects on landscape health (17). Average per acre cost to growers to implement food safety modifications to meet the "Leafy Green Marketing Agreement" (LGMA) was $13.60 based on a survey conducted in 2008 and 2009 (18).

Sources of preharvest contamination with foodborne pathogens can occur from a variety of sources (e.g., irrigation and run-off water, soil amendments, such as manure, fecal deposition from intruding domesticated and wild animals). In addition, management practices (e.g., worker hygiene, buffer zones) and geospatial factors (e.g., soil characteristics) can significantly modulate the risk of contamination from different sources (2, 3, 12, 19-21). A number of studies have shown that water can act as both a source of pathogens and vehicle of pathogen introduction to preharvest environments and produce (20, 22-25). For example, surface water has been reported to have a wide range of *Salmonella* (6% to 80%) and *L. monocytogenes* (6.4% to 62%) prevalence (24, 26-29). In particular, *Salmonella* prevalence of 6-9% has been reported for water samples obtained from produce growing regions in California (CA) and New York State

(NYS) (14, 23). Manas et al. (30) determined that lettuce plants irrigated with non-potable water had significantly higher rates of total coliforms and *Salmonella* contamination than lettuce irrigated with drinking water. A number of studies also have linked sporadic or repeated contamination events in produce fields to wildlife fecal deposits (21) with a variety of bacterial foodborne pathogens, including *Salmonella* and *L. monocytogenes* regularly isolated from fecal samples collected from wildlife and domesticated animals (3, 31-36). *Salmonella* can also survive in the soil for long periods of time (e.g., up to 230 days in one study (37)) when introduced by contaminated poultry or cow manure. A study of farm management practices in Minnesota and Wisconsin found that the use of manure significantly increased the risk of *E. coli* contamination in organic (OR 13.2) and semi-organic (OR 12.9) produce (38). Another study demonstrated that worker hygiene (e.g., portable toilets, hand-washing stations) and trainings were important in reducing the likelihood of generic *E. coli* contamination at the preharvest level (39).

While a number of studies (3, 12, 23, 38-41) have suggested that specific farm management practices may impact pathogen contamination in the preharvest environment, we are not aware of any studies that used statistical methods to quantitatively assess the risk of pathogen contamination associated with specific field-level management practices. These types of data are essential to allow for identification of practices that can significantly increase or decrease the likelihood of field-level contamination, in order to facilitate implementation of science-based preventive controls. Thus, the purpose of this study was to (i) evaluate the prevalence of *Salmonella* and *L. monocytogenes* isolated from environmental samples (soil, drag swab, and water) and (ii) identify field-level management practices associated with presence of *Salmonella* or *L. monocytogenes*.

MATERIALS AND METHODS

Study Design. Twenty-one produce farms in NYS were enrolled in a cross-sectional study. Enrollment was based on the willingness of the grower to participate in the study. Participation entailed permission to collect environmental samples from produce fields on the farm and agreement to fill out a questionnaire regarding field-level management practices associated with each field that was sampled. Farms were located in three regions of NYS with five in western New York, 12 in central New York and four in eastern New York. Farm visits were performed over a five-week period in June and July 2012. At least ten fields were selected per farm. A single composite soil sample (consisting of five subsamples of soil from five locations in the field) and an area drag swab sample were collected for each field (using a sampling area of approximately 0.2 ha). Additionally, samples were collected from water sources that were (i) used for field irrigation (n=23) or (ii) not used for field irrigation, but were within 50 m from a sampled field (n=51). Six hundred samples were collected for the study (263 composite soil samples, 263 area drag swab samples and 74 water samples).

Questionnaire Design. A questionnaire was developed to obtain data on field-level practices identified in literature as possible factors (e.g., manure application, irrigation water) that influence the risk of preharvest contamination. The interview form included questions to (i) obtain general farm characteristics (15 questions) and (ii) information on sampled fields (11 field questions). Seven of the 11 field-specific questions were time-dependent. For instance, growers were asked the last time a sampled field was irrigated, with answer options: within 3 d, 4-7 d, 8-14 d, and over 14 d/never. One of the time-dependent questions (frequency of irrigation) had two follow-up questions. The two follow-up questions were (i) source of irrigation water (e.g., pond) and (ii) type of irrigation system used (e.g., drip). The remaining four specific field questions were not time-dependent. For example, growers were asked if the field had a buffer zone (i.e.,

defined as at least a 5 m strip where no produce was grown). Questionnaires were administered by a single interviewer (LS) and completed at the time of sample collection in a face-to-face interview, which lasted approximately 1 h. Data were coded from the questionnaires, entered into Excel (Microsoft, Redmond, WA) and imported into SAS 9.3 (SAS Institute Inc., Cary, NC).

Sample Collection. Samples were collected as previously detailed by Strawn et al. (14). Briefly, latex gloves and disposable plastic boot covers (Nasco, Fort Atkinson, WI) disinfected with 70% ethanol were worn and changed for sample collection at each field. Five soil samples per field were taken using sterile scoops (Fisher Scientific, Hampton, NH) at least six inches (15.2 cm) below the surface (sub-surface soil) and deposited in separate sterile Whirl-Pak bags (Nasco). A pre-moistened drag swab (30 mL of buffered peptone water (BPW; Becton Dickinson, Franklin Lakes, NJ) in a sterile Whirl-Pak bag), as previously described by Uesugi et al. (42), was dragged through the field (side to side in 10-m increments, around perimeter of field) for 10 min. Water samples (n=74) were collected directly into sterile 250 mL jars; a sampling pole (Nasco) was used if necessary (i.e., for creeks, ponds). Surface water samples were taken a minimum of 2 m from the water's edge and 0.3 m below the surface. All samples were transported on ice, stored at 4±2°C and processed within 24 h of collection.

Sample Preparation. Samples were prepared for two enrichment schemes to allow for separate isolation and detection of *Salmonella* and *L. monocytogenes*. Composite soil samples were prepared by combining 5 g portions of each of the five subsamples of soil collected in a field in duplicate. Both 25 g composite soil samples were deposited into sterile filter Whirl-Pak bags. Individual drag swab samples were combined with BPW and hand massaged for 2 min, squeezed, and 10 mL of the liquid contents aseptically transferred to each of two sterile filter Whirl-Pak bags. Water samples were tested according to Environmental Protection Agency (EPA) standard methods (43, 44). Briefly, each water sample collected (250 mL) was passed

through a 0.45 μm filter unit (Nalgene, Rochester, NY). The subsequent filter was then aseptically removed and cut in half, and each portion transferred to a separate sterile filter Whirl-Pak bag.

***Salmonella* and *L. monocytogenes* Detection and Isolation.** *Salmonella* (45) and *L. monocytogenes* (46) detection and isolation was performed using modified versions of the procedures outlined in the US Food and Drug Administration's Bacteriological Analytical Manual (FDA BAM). No quantification of *Salmonella* or *L. monocytogenes* was performed. Briefly, for *Salmonella* detection and isolation, samples were diluted 1:10 with tryptic soy broth (TSB; Becton Dickinson) and allowed to stand for 2 h at room temperature (23±2°C). After incubation at 35±2°C for an additional 24 h, two aliquots (1.0 and 0.1 mL) were transferred to 9 and 9.9 mL of tetrathionate (TT; Oxoid; Cambridge, UK) and Rappaport Vassiliadis (RV; Oxoid) broths, respectively. Both selective enrichment broths were incubated at 42°C in a shaking water bath for 24 h. A 50 μl aliquot of TT and RV broths were plated onto xylose lysine deoxycholate agar (XLD; Neogen, Lansing, MI) and CHROMagar *Salmonella* (CHROMagar Company; Paris, France), and incubated at 35 and 37±2°C for 24 and 48 h, respectively. Up to four presumptive *Salmonella* colonies per selective enrichment and plating medium combination (e.g., TT-XLD, RV-XLD) were sub-streaked to brain heart infusion agar (BHI; Becton Dickinson) and incubated at 37±2°C for 24 h. Presumptive *Salmonella* colonies were confirmed by a polymerase chain reaction (PCR) assay that detects the gene, *invA* (47). For *L. monocytogenes*, all samples were diluted 1:10 with buffered *Listeria* enrichment broth (BLEB; Becton Dickinson) and incubated at 30±2°C for 24 h. *Listeria* Selective Enrichment Supplement (Oxoid) was added to enrichments at 4 h. At 24 and 48 h, 50 μl of enrichment was streaked onto modified Oxford agar (MOX, Becton Dickinson) and *L. monocytogenes* plating medium (LMPM, Biosynth International, Itasca, IL). MOX and LMPM plates were incubated for 48 h at

30 and 35±2°C, respectively. Up to four *L. monocytogenes* presumptive colonies per plating medium and time combination (e.g., MOX 24 h or LMPM 48 h) were sub-streaked onto BHI. BHI agar plates were incubated for 37±2°C for 24 h. Presumptive *L. monocytogenes* colonies were confirmed by PCR amplification and sequencing of the partial *sigB* gene (48-50). Controls were processed in parallel with each pathogen detection and isolation scheme. *Salmonella* ATCC 700408 (FSL F6-826; (51) and *L. monocytogenes* FSL R3-001 (52), were used as positive controls. Sterile enrichment media were used as negative controls.

Classification of Isolates. There were four isolation schemes for each *Salmonella* (TT-XLD, RV-XLD, TT-CHROME, and RV-CHROME) and *L. monocytogenes* (LMPM 24 h, MOX 24 h, LMPM 48 h, and MOX 48 h); one isolate from each "isolation scheme" was used for subtyping (as detailed below), yielding up to four representative isolates per pathogen-positive. Representative isolates were streaked from frozen culture onto BHI and incubated at 37°C for 18 h, and a well-isolated colony selected. *Salmonella* serotyping using the White-Kauffman-Le Minor scheme (53) was conducted by the Wadsworth Center, New York State Department of Health (Albany, NY). Nucleotide sequences of *sigB* from *L. monocytogenes* isolates were obtained by Sanger sequencing performed by the Cornell University Life Sciences Core Laboratories Center. Allelic types (AT), as defined by a unique combination of polymorphisms (54, 55), were assigned by comparison of *sigB* sequences to an internal reference database.

Statistical Analysis. Separate statistical analyses for *Salmonella* and *L. monocytogenes* were performed in SAS 9.3. An initial descriptive analysis was performed to calculate *Salmonella* and *L. monocytogenes* prevalence for all samples (n=600) and each sample type collected: soil (n=263), drag swab (n=263) and water (n=74). Univariate associations between pathogen-positive terrestrial samples and region and week sampled were determined using a chi-square test or Fisher's exact test (if the expected frequency in any cell was less than 5). Confidence intervals

(95%) were calculated assuming a binominal distribution. Individual *p* values are reported for each test.

A field was used as the unit of analysis for model development to identify field-level risk factors associated with *Salmonella* and *L. monocytogenes* contamination in produce fields. A field was considered positive if either a soil or drag swab sample collected from that field was confirmed culture positive for the respective pathogen. Chi-square or Fisher's exact tests were computed for each of the 11 specific field questions (i.e., factors). Factors determined to be significant ($P \leq 0.05$) were retained as candidates for subsequent multivariate analysis. General linear mixed models (GLIMMIX) procedure was used to model the association between each candidate factor (univariate analysis) or factors (multivariate analysis) and the outcome (*Salmonella* or *L. monocytogenes*-positive/negative field). Fields within farm were not independent; therefore, farm was included in the model (as a random effect). Effect estimates (β), standard errors (SE), odds ratios (OR), 95% confidence intervals (CI), and *P* values were determined for each candidate factor. Potential collinearity among the candidate factors was evaluated by Spearman's rank correlation coefficient test. Multivariable models were built using a stepwise selection method and accessed by fit statistics, such as Akaike's information criteria and Schwarz' Bayesian criterion. The final model retained only variables that significantly improved the fit of the model ($P \leq 0.05$). Interaction terms were also tested, but none were significant.

Isolate Storage and Data Access. All isolates were preserved at -80°C in 15% glycerol. Isolate information and subtyping data from this study are archived and available through the Food Microbe Tracker database (http://www.foodmicrobetracker.com).

RESULTS

***Salmonella* and *L. monocytogenes* Prevalence in Terrestrial Samples**. The prevalence of *Salmonella* in terrestrial samples (n=263 soil and n=263 drag swab) was 3.4% (18/526). *Salmonella* prevalence was higher among soil samples (13/263), compared to drag swab samples (5/263). *Salmonella* was detected in 6.1% of fields sampled (16/263). For two fields, both soil and drag swab samples were positive for *Salmonella*. No significant difference was observed in the *Salmonella* prevalence in soil and drag swab samples by region (P = 0.4 and 0.9, respectively) and week sampled (P = 0.9 and 0.6, respectively). Furthermore, no significant difference was observed for the field-level prevalence of *Salmonella* by region (P = 0.8) and week sampled (P = 0.9).

The prevalence of *L. monocytogenes* in terrestrial samples (n=263 soil and n=263 drag swab) was 9.7% (51/526). *L. monocytogenes* prevalence in soil and drag swab samples was 11% (30/263) and 8% (21/263), respectively. *L. monocytogenes* was detected in 46 of the 263 fields sampled (17.5%). Five fields had both soil and drag swab samples that were positive for *L. monocytogenes*. No significant difference was found in the *L. monocytogenes* prevalence in soil and drag swab samples by region (P = 0.2 and 0.3, respectively) and week sampled (P = 0.7 and 0.2, respectively). In addition, no significant difference was found for the field-level prevalence of *L. monocytogenes* by region (P = 0.9) and week sampled (P = 0.1).

***Salmonella* and *L. monocytogenes* Prevalence in Water Samples.** The prevalence of *Salmonella* and *L. monocytogenes* in water samples was 11% (8/74) and 30% (22/74), respectively. Samples were collected from irrigation (n = 23) and non-irrigation water sources (within 50 m of a sampled field; n=51) (Table 3.1).

The prevalence of *Salmonella* and *L. monocytogenes* in water samples used for irrigation was 4% (1/23) and 9% (2/23), respectively. Fourteen of the samples collected from irrigation

sources were obtained from engineered water sources (e.g., well, municipal), which were of potable water quality; all of the samples were negative for *Salmonella* and *L. monocytogenes*. The remaining nine water samples were from surface water sources (1 creek and 8 pond samples); three samples from ponds used for field irrigation tested positive for *Salmonella* (1 sample) and *L. monocytogenes* (two samples). All fields sampled using these irrigation water sources were negative for the presence of *Salmonella* and *L. monocytogenes* (Table 3.1).

Salmonella and *L. monocytogenes* were detected in 14% (7/51) and 39% (20/51), respectively, of water samples obtained from non-irrigation sources and within 50 m of a sampled field. Water samples were collected from three source types: ponds (n=17), roadside or field buffer ditches (n=13), and flowing surface water (e.g., rivers, creeks, or streams (n=21)). The prevalence of *Salmonella* was higher in roadside or field buffer ditch samples (23%, 3/13), compared to pond (12%, 2/17) and flowing surface water (10%, 2/21) samples. The prevalence of *L. monocytogenes* was highest in pond samples (59%, 10/17), compared to roadside or field buffer ditch (39%, 5/13) and flowing surface water (24%, 5/21) samples (Table 3.1).

Table 3.1 *Salmonella* and *L. monocytogenes* prevalence in water samples collected from irrigation and non-irrigation water sources

		Prevalence (Frequency) of	
Category	Count	*Salmonella*	*L. monocytogenes*
Not used for irrigation[a]	51	14 (7)	39 (20)
Pond	17	12 (2)	59 (10)
Ditch[b]	13	23 (3)	39 (5)
River/creek/stream	21	10 (2)	24 (5)
Used for irrigation	23	4 (1)	9 (2)
Engineered[c]	14	0 (0)	0 (0)
Pond	8	13 (1)	25 (2)
River/creek/stream	1	0 (0)	0 (0)
Total	74	11 (8)	30 (22)

[a] Water samples not used for irrigation were collected within 50 m from a sampled field.

[b] Ditch was defined as either a roadside ditch (located between road and field) or a runoff ditch (located between landscape feature (e.g., a pasture) and field; often part of a buffer zone).

[c] Engineered water was defined as water from a well or municipal source (i.e., a potable water source).

Characterization of *Salmonella* and *L. monocytogenes* Isolated from Terrestrial and Water Samples. Serotyping was performed on one representative *Salmonella* isolate per isolation scheme, which yielded 35 *Salmonella* isolates from the 26 positive samples. Three of the 26 samples yielded isolates with more than one serotype. *Salmonella* Give and Typhimurium were isolated from a single water sample (isolation schemes TT-Chrome and TT-XLD, respectively), *Salmonella* Agona and Tennessee were isolated from a drag swab sample (isolation schemes RV-Chrome and RV-XLD) and *Salmonella* Senftenburg and Newport were isolated from a soil sample (isolation schemes RV-Chrome and RV-XLD, respectively). The remaining 23 *Salmonella*-positive samples represented one serotype. These isolates were identified as *Salmonella* serotypes Newport (8 samples), Cerro (5 samples), Thompson (5 samples), Agona (2 samples), IV 40:z4,z32:- (2 samples), and Give (1 sample). For the two fields where *Salmonella* was isolated in both soil and drag swab samples, the same serotype (Cerro) was isolated in both sample types from one field, whereas different serotypes (Thompson and Cerro) were isolated in the soil and drag swab samples from the other field.

Two-hundred and sixteen *L. monocytogenes* isolates (one isolate per isolation scheme) were subtyped based on alignment of *sigB* nucleotide sequences. None of the four isolation schemes yielded different subtypes for any sample. The 73 representative *L. monocytogenes* isolates (from the 73 *L. monocytogenes*-positive samples) yielded nine different allelic types that represented *L. monocytogenes* lineage I (29 isolates, 5 ATs), II (41 isolates, 3 Ats), and IIIa (3 isolates, 1 AT). *L. monocytogenes* was detected in both soil and drag swab samples for five fields. The same subtype was identified in soil and drag swab samples in two fields (AT 57 and AT 59), whereas different subtypes (Ats 57 and 61, Ats 78 and 137, and Ats 57 and 58) were isolated in the soil and drag swab samples from three fields.

Risk Factors Associated with *Salmonella* Contamination of Produce Fields. Three of the 11 field management practices evaluated were significantly associated with a *Salmonella*-positive field by univariate analysis (manure application, soil cultivation, and buffer zone; Table 3.2). Fields where manure was applied within a year prior to sample collection had higher odds of *Salmonella* isolation (OR = 19.0, 95% CI = 4.9, 77.0), compared with fields where manure had not been applied. Fields where soil was cultivated within 7 d prior to sample collection were approximately 6 times more likely (OR = 6.3, 95% CI = 1.6, 23.0) to be *Salmonella*-positive, compared with fields where soil was not cultivated for at least 30 d. The presence of a buffer zone was shown to have a protective effect and reduced the likelihood of a *Salmonella*-positive field by 5 times (OR = 0.2, 95% CI = 0.1, 0.5; Table 3.2).

Table 3.2 Univariate analyses of management practices that influence the likelihood of *Salmonella* being detected in a produce field (based on testing of soil and drag swab samples)

Factor	Description	β-coefficient	SE[a]	OR[b]	95% CI[c]	*P* value
Manure	Last time manure was applied to field					
	1 = within 365 d	3.0	0.7	19	4.9, 77	<0.001
	2 = over 365 d	0.4	0.9	1.5	0.2, 9.4	0.681
	3 = not been applied	0	-	1.0	-	
Soil Cultivation	Last time soil in field was cultivated					
	1 = within 7 d	1.8	0.7	6.3	1.6, 23	0.008
	2 = 8 to 14 d	0.5	0.9	1.6	0.3, 9.9	0.625
	3 = 15 to 30 d	-0.9	1.2	0.4	0.1, 4.2	0.461
	4 = over 30 d	0	-	1.0	-	
Buffer zone[d]	Does field have a buffer zone					
	1 = yes	-1.7	0.5	0.2	0.1, 0.5	0.002
	2 = no	0	-	1.0	-	

[a] standard error
[b] odds ratio
[c] confidence interval
[d] buffer zone was defined as a strip of land where no produce was grown, approximately 5 m wide

Examination of Spearman's rank correlation coefficients for the three retained candidate factors from the univariate analysis showed a correlation between application of manure and soil cultivation of a field. Therefore, three multivariable models were evaluated: model 1 = manure application, soil cultivation, and buffer zone; model 2 = manure application and buffer zone; and model 3 = soil cultivation and buffer zone. In the multivariate model with the best fit (i.e., model 2; Table 3.3), application of manure to a field within a year prior to sample collection was associated with a higher likelihood of *Salmonella* being detected in a field (OR = 16.7, 95% CI = 3.0, 94.4) as compared with fields where manure had not been applied. Presence of a buffer zone was associated with a lower likelihood of *Salmonella* being detected in a field (OR = 0.1, 95% CI = 0.03, 0.6), as compared to absence of a buffer zone (Table 3.3).

Table 3.3 Multivariate final model[a] of risk factors that influence the likelihood of *Salmonella* being detected in a produce field (based on testing of soil and drag swab samples).

Factor	Description	β-coefficient	SE[b]	OR[c]	95% CI[d]	*P* value
Manure	Last time manure was applied to field					
	1 = within 365 d	2.8	0.9	16.7	3.0, 94.4	0.002
	2 = over 365 d	0.3	1.1	1.3	0.2, 11	0.789
	3 = not been applied	0	-	1.0	-	
Buffer zone[e]	Does field have a buffer zone					
	1 = yes	-2.0	0.7	0.1	0.03, 0.6	0.008
	2 = no	0	-	1.0	-	
Farm	Random effect	1.6	1.5			

[a] final model developed in PROC GLIMMIX; farm as random effect.
[b] standard error
[c] odds ratio
[d] confidence interval
[e] buffer zone was defined as a strip of land where no produce was grown, approximately 5 m wide

Risk Factors Associated with *L. monocytogenes* Contamination in Produce Fields. Six of the 11 field management practices were significantly associated with a *L. monocytogenes*-positive field by univariate analysis (manure application, reporting of wildlife, worker activity, irrigation, soil cultivation and reporting of a buffer zone; Table 3.4); five of these six factors were time-dependent. Fields where manure was applied within a year prior to sample collection had 7 times higher odds of *L. monocytogenes* isolation (OR = 7.0, 95% CI = 3.1, 15.4), compared with fields where manure had not been applied. Fields where growers reported observation of wildlife within 3 d prior to sample collection had higher odds of *L. monocytogenes* isolation (OR = 4.4, 95% CI = 1.2, 15.6), compared with fields where growers did not report observation of wildlife for at least 7 d. Fields where soil was cultivated within 7 d prior to sample collection were approximately 8 times more likely (OR = 8.1, 95% CI = 3.3, 19.6) to be *L. monocytogenes*-positive, compared with fields where soil was not cultivated for at least 30 d. Fields with recent worker activity (within 3 d prior to sample collection) had 10.5 times higher odds of *L. monocytogenes* isolation (OR = 10.5, 95% CI = 2.3, 47.5), compared with fields where workers had been absent for longer than 30 d. A number of other worker related factors did not show significant associations with *L. monocytogenes* contamination, including delivery of food safety training (in the native language), presence of portable toilets and hand washing stations (within a quarter mile of fields), cleaning frequency of toilets, and posting of signs advocating food safety and or sanitation best practices in changing areas; for most of these factors a high level of compliance with "best practices" was reported (e.g., all farms reported cleaning toilets at least once a week). Fields irrigated within 3 d prior to sample collection had nearly 5.5 times higher odds of *L. monocytogenes* isolation (OR = 5.3, 95% CI = 2.4, 12.0), compared with fields irrigated at least 14 d ago. Furthermore, no significant difference was observed in *L. monocytogenes*-positive fields for irrigation type (overhead versus drip). Lastly, presence of a

buffer zone was shown to have a protective effect and reduced the likelihood of a *L. monocytogenes*-positive field (OR = 0.5, 95% CI = 0.2, 0.9; Table 3.4).

Table 3.4 Univariate analyses of management practices that influence the likelihood of *L. monocytogenes* being detected in a produce field (based on testing of soil and drag swab samples)

Factor	Description	β-coefficient	SE[a]	OR[b]	95% CI[c]	*P* value
Manure	Last time manure was applied to field					
	1 = within 365 d	1.9	0.4	7.0	3.1, 15.4	<0.001
	2 = over 365 d	-0.4	0.5	0.6	0.2, 1.7	0.381
	3 = not been applied	0	-	1.0	-	
Wildlife[d]	Last time wildlife was observed in field					
	1 = within 3 d	1.5	0.6	4.4	1.2, 15.6	0.022
	2 = 4 to 7 d	-0.2	0.7	0.8	0.2, 3.1	0.725
	3 = 8 to 30 d	0	-	1.0	-	
Worker Activity	Last time workers were in the field					
	1 = within 3 d	2.4	0.8	10.5	2.3, 47.5	0.003
	2 = 4 to 7 d	0.7	0.8	1.9	0.4, 9.9	0.439
	3 = 8 to 30 d	1.0	0.9	2.6	0.5, 14.7	0.281
	4 = over 30 d	0	-	1.0	-	
Irrigation	Last time field was irrigated					
	1 = within 3 d	1.7	0.4	5.3	2.4, 12.0	<0.001
	2 = 4 to 7 d	-0.3	0.5	0.8	0.3, 2.2	0.599
	3 = 8 to 14 d	-1.3	0.7	0.3	0.1, 1.1	0.061
	4 = over 14 d/not irrigated	0	-	1.0	-	
Soil Cultivation	Last time soil in field was cultivated					
	1 = within 7 d	2.1	0.5	8.1	3.3, 19.6	<0.001
	2 = 8 to 14 d	0.6	0.6	1.8	0.6, 5.7	0.293
	3 = 15 to 30 d	-0.4	0.6	0.7	0.2, 2.2	0.540
	4 = over 30 d	0	-	1.0	-	
Buffer zone[f]	Does field have a buffer zone					
	1 = yes	-0.8	0.4	0.5	0.2, 0.9	0.049
	2 = no	0	-	1.0	-	

[a] standard error
[b] odds ratio
[c] confidence interval
[d] answer option 4 (never) was not selected in the questionnaire; therefore it was excluded from analysis
[e] a worker constituted a man or woman in the field, not in the cab of farm equipment (e.g., tractor)
[f] buffer zone was defined as a strip of land where no produce was grown, approximately 5 m wide

Correlation was evaluated between the six factors retained by univariate analysis using Spearman's rank correlation coefficients. Similar to the findings for *Salmonella*, a correlation was observed between manure application and soil cultivation of a field. The three multivariable models evaluated were these (i) model 1 = manure application, reported observation of wildlife, worker activity, irrigation, soil cultivation, and buffer zone; (ii) model 2 = manure application, reported observation of wildlife, worker activity, irrigation, and buffer zone; and (iii) model 3 = reported observation of wildlife, worker activity, irrigation, soil cultivation, and buffer zone. The multivariate model with the best fit was model 3 (Table 3.5). In this model, reported observation of wildlife in a field (OR = 6.1, 95% CI = 1.3, 28.4) and irrigation of a field (OR = 6.0, 95% CI = 2.0, 18.1) within 3 d prior to sample collection were associated with higher odds of *L. monocytogenes* isolation. Fields where soil was cultivated within 7 d prior to sample collection were nearly 3 times more likely to be *L. monocytogenes*-positive, compared with fields where soil was cultivated at least 30 d ago (OR = 2.9, 95% CI = 1.1, 8.6; Table 3.5).

Table 3.5 Multivariate final model[a] of risk factors that influence the likelihood of *L. monocytogenes* being detected in a produce field (based on testing of soil and drag swab samples)

Factor	Description	β-coefficient	SE[b]	OR[c]	95% CI[d]	*P* value
Wildlife[e]	Last time wildlife was observed in field					
	1 = within 3 d	1.8	0.8	6.1	1.3, 28.4	0.021
	2 = 4 to 7 d	-0.02	0.8	1.0	0.2, 4.8	0.978
	3 = 8 to 30 d	0	-	1.0	-	
Irrigation	Last time field was irrigated					
	1 = within 3 d	1.8	0.6	6.0	2.0, 18.1	0.001
	2 = 4 to 7 d	0.2	0.7	1.2	0.3, 4.5	0.793
	3 = 8 to 14 d	-0.8	0.8	0.4	0.1, 2.0	0.288
	4 = over 14 d/not irrigated	0	-	1.0	-	
Soil Cultivation	Last time soil in field was cultivated					
	1 = within 7 d	1.1	0.6	2.9	1.1, 8.6	0.050
	2 = 8 to 14 d	0.3	0.7	1.4	0.4, 5.1	0.660
	3 = 15 to 30 d	-0.9	0.7	0.4	0.1, 1.7	0.224
	4 = over 30 d	0	-	1.0	-	
Farm	Random effect	0.1	3.3			

[a] final model developed in PROC GLIMMIX; farm as random effect
[b] standard error
[c] odds ratio
[d] confidence interval
[e] answer option 4 (never) was not selected in the questionnaire; therefore it was excluded from analysis

DISCUSSION

Our study reported here is one of the first to quantitatively identify management practices that are associated with an increased or decreased likelihood of *Salmonella* and *L. monocytogenes* isolation in produce fields. In a univariate analysis, six factors (manure application, reported observation of wildlife, worker activity, irrigation, soil cultivation, and buffer zone presence) were identified as significant risk factors for *Salmonella* or *L. monocytogenes* contamination. Five of the six risk factors were influenced by time of application to fields, suggesting that adjustments to current practices may reduce the potential for produce contamination with minimal costs to growers.

Some Risk Factors Influence the Likelihood of both *Salmonella* and *L. monocytogenes* Isolation in Fields. Based on the separate univariate analysis of *Salmonella* and *L. monocytogenes* data, we identified three risk factors that significantly affected the likelihood of both *Salmonella* and *L. monocytogenes* detection. As adjustments of management factors related to these risk factors have the potential to reduce contamination with both of these key pathogens, these three factors are discussed below.

Our data specifically showed that recent cultivation of fields (i.e., within 7 days of sample collection) was significantly associated with an increased likelihood of both *Salmonella* and *L. monocytogenes* isolation from fields. Soil cultivation was also found to be a significant risk factor in the final multivariate model for *L. monocytogenes* isolation. A likely explanation for these findings is that pathogens present in the sub-surface soil are exposed to the surface when soil is cultivated, making them more likely to be detected, and possibly also more likely to contaminate produce. Furthermore, the likelihood of pathogen isolation will decrease over time after cultivation, due to exposure to environmental conditions (e.g., UV light) that reduce pathogen loads. This model is supported by previous studies (14, 23, 42) that have shown the

presence and persistence of *Salmonella* and *L. monocytogenes* in sub-surface soil. For example, *Salmonella* was detected in 2.6% and 2.0% of soil samples collected from produce growing regions in CA and NYS, respectively; while *L. monocytogenes* prevalence in soil was 9% in NYS preharvest environments (14, 23). Interestingly, Park et al. (39) observed that spinach contamination with generic *E. coli* was less likely when a field was cultivated prior to the growing season; this may reflect that cultivation at time points considerably before sampling (e.g., > 7 d before) will reduce overall pathogen loads by exposing pathogens present in the sub-surface soil to UV and other inactivating conditions (e.g., desiccation). This hypothesis is supported by the observation that *Salmonella* and *Listeria* numbers in inoculated livestock waste declined more rapidly when this material was spread on the surface of soil as compared with incorporation into the soil where it would be protected from exposure to environmental conditions (e.g., UV light, harsh temperatures) (56).

Application of manure was also identified as a significant factor that increased the odds of both *Salmonella* and *L. monocytogenes* isolation in fields. Numerous studies (37, 39, 41, 57-64) have demonstrated that the application of manure to soils can introduce pathogens and may facilitate long term persistence of pathogens in soil. One study observed *Salmonella* to persist in manure, manure-amended non-sterilized soil, and manure-amended sterilized soil for 184, 332, and 405 days, respectively (60). However, the association between pathogen contamination of fields and manure application has not been previously described using commercial produce farms. Some studies (38, 39) have investigated the association between generic *E. coli* contamination of preharvest produce samples and application of manure to fields. One study observed that generic *E. coli* contamination was lower in spinach samples collected over a two year period if the application of manure occurred greater than 200 d prior to sample collection (39), while another study observed that *E. coli* prevalence in preharvest produce samples

collected was not affected by the application of manure between 90 to 120 d prior to sampling (38). Our results suggest that application of manure to fields can significantly influence the risk of both *Salmonella* and *L. monocytogenes* contamination; therefore, management of manure before application is essential. Manure management practices, such as aging, treating, and handling of manure before application have been shown to affect the survival of foodborne pathogens in manure (38, 56, 65). For example, one study (65) showed that composting cow manure before application was effective at killing *Salmonella*, supporting that management of manure before application to fields may limit or reduce the risks associated with manure use in produce preharvest environments.

In addition, the likelihood of *Salmonella* and *L. monocytogenes* isolation in fields was significantly decreased if growers reported presence of a buffer zone, defined as a zone of at least 5 m separating the edge of produce fields from potential environmental pathogen reservoirs (e.g., forests, roads, waterways, livestock operations). These data suggest that even buffer zones narrower than the 10 m (30 ft.) recommended in the 2012 version of the LGMA (table 6 (66)) are associated with reduced pathogen prevalence. Surprisingly, there is little science-based research to support the hypothesis that presence of a buffer zone is associated with decreased pathogen prevalence in preharvest environments. Therefore, in our study we formally tested the hypothesis that presence of a buffer zone is associated with decreased pathogen prevalence (i.e., *Salmonella* and *L. monocytogenes*) in produce fields. Some previous studies (67-69) suggest that vegetative buffer zones may be effective in reducing bacterial pathogen loads in sewage runoff and wastewater from animal facilities. Vegetative buffer zones and non-agricultural lands adjacent to produce fields (e.g., riparian, wetlands, grasslands) also offer a variety of ecological benefits (16, 17, 69, 70). Combined these data suggest that the effects, on pathogen prevalence, of buffer zones and non-agricultural lands adjacent to produce fields may be driven by complex ecological

interactions that will require further field studies that include mathematical modeling efforts. These research efforts will also need to define the effects of different types of buffer zones (i.e., bare strips, specific vegetation) and the quantitative relationship between buffer zone width and type, and pathogen reduction.

Some Risk Factors Specifically Increase the Likelihood of *L. monocytogenes* Isolation in Fields. While some risk factors increased or reduced the likelihood of both *Salmonella* and *L. monocytogenes*, others (worker activity, reported wildlife observation, and irrigation) were solely identified to increase the likelihood of *L. monocytogenes* detection in fields. Worker activity was significantly associated with an increased likelihood of *L. monocytogenes* isolation in fields by univariate analysis, but was not significant in the multivariate analysis. However, reported observation of wildlife and irrigation of fields were significantly associated with higher odds of *L. monocytogenes* isolation by multivariate analysis and are discussed below.

Reported observation of wildlife was based on visual confirmation (i.e., sighting of wildlife in a field) by the grower or his/her staff (e.g., field supervisor). We acknowledge that growers who have their farms and food safety programs (e.g., GAPs) frequently audited may be less inclined to report presence of wildlife because they are aware of the risks associated with wildlife in fields. While growers who have their farms and food safety programs infrequently audited may be more forthcoming to report presence of wildlife. Future studies may choose to measure the impact of wildlife and potential pathogen contamination by objective measures (e.g., the use of infrared cameras to detect wildlife in fields). Our study does provide quantitative data to support previous studies (2, 12, 19, 25, 32) that suggested that wildlife may be a source of pathogen contamination in fields. Furthermore, wildlife has also been suspected as the source of pathogen contamination in a number of produce-associated outbreaks (13, 70, 71). While reported observation of wildlife was shown to be a risk factor increasing the likelihood of *L.*

monocytognes isolation in fields, this finding may be site specific to NYS, or parts of NYS; Langholz and Jay-Russell have discussed that pathogen prevalence in wildlife may be dependent on geographic location and local landscape characteristics (70).

Recent irrigation was also shown to significantly increase the odds of *L. monocytogenes* isolation in fields. Water has been identified as a major reservoir for pathogens and irrigation a vehicle for transmission of pathogens to fields and produce (12, 30, 41, 72-75). *L. monocytogenes* is often found in various water sources with prevalence reported from <1% to 29% (14, 76, 77). We also observed, here, a high prevalence of *L. monocytogenes* in water, particularly surface water sources (e.g., ponds). Steele and Odumeru (72) observed surface water had the most variable microbial quality, and if contaminated, could lead to widespread contamination of crops. Our findings suggest that detection of *L. monocytogenes* in fields was only more likely if irrigation occurred within a couple of days prior to sample collection. Two studies have also shown an association between pathogen detection and time of irrigation or water application. One study observed that *Salmonella* when sprayed on tomatoes was not able to be recovered from the tomatoes after two days (78). The second study observed that the risk of *E. coli* contamination in spinach samples decreased when irrigation in a field occurred >5 d prior to sample collection. In addition to *L. monocytogenes* introduction with irrigation water, the association of irrigation with an increased frequency of *L. monocytogenes* detection may also reflect the fact that moist soils may facilitate *L. monocytogenes* growth or detection, consistent with previous studies that reported a higher *L. monocytogenes* prevalence in moist soils (14, 79). Overall, our data suggest that avoiding irrigation at least 3 d before harvest (if possible and feasible) may reduce potential *L. monocytogenes* contamination to produce and possibly the transfer into packinghouses, from soil in the fields.

Conclusions. This study provides quantitative data on management practices that represent

potential risk factors for produce field contamination. A majority of research previously conducted to investigate these risk factors has been pathogen inoculation-based or targeted presence of indicator organisms (i.e., generic *E. coli*). Such studies are commonly employed because the prevalence of foodborne pathogens (*Salmonella*, Shiga toxin producing *E. coli*) in produce production environments is low. Statistically robust analyses are difficult to conduct unless a sufficient number of pathogen-positive samples are obtained, and this generally requires an extremely large sample size. Large sample sizes in environmental field studies are often difficult to achieve due to considerable labor and financial costs, and difficulties gaining access to commercial operations. We focused on only eleven key management practices previously discussed as risk factors for preharvest contamination, limited the number of levels within each factor, and opted for a statistical procedure to deal with farm as a confounder; in order, to prevent bias and misinterpretation of results (e.g., spurious relationships). This study was conducted in NYS, thus risk factors identified may not always be appropriate in other produce growing regions in the US or elsewhere. Additionally, fields were sampled over a five-week period in June and July, as a result risk factors identified may not be applicable to other time periods (e.g., late in the growing season). Despite some limitations, this study is one of the first to use field collected data to provide quantitative data on management practices associated with detection of *Salmonella* and *L. monocytogenes* (two foodborne pathogens of concern to the produce industry). These findings will assist growers in (i) evaluating their current on-farm food safety plans (e.g., GAPs), (ii) implementing preventive controls that reduce the risk of preharvest contamination, and (iii) making more informed decisions related to field practices prior to harvest.

.

REFERENCES

1. **Painter JA, Hoekstra RM, Ayers T, Tauxe RV, Braden CR, Angulo FJ, Griffin PM.** 2013. Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998-2008. Emerg. Infect. Dis. **19:**407-415.
2. **Doyle MP, Erickson MC.** 2008. Summer meeting 2007 – the problems with fresh produce: an overview. J. Appl. Microbiol. **105:**317-330.
3. **Hanning IB, Nutt JD, Ricke SC.** 2009. Salmonellosis Outbreaks in the United States Due to Fresh Produce: Sources and Potential Intervention Measures. Foodborne Pathog. Dis. **6:**635-648.
4. **Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM.** 2011. Foodborne Illness Acquired in the United States-Major Pathogens. Emerg. Infect. Dis. **17:**7-15.
5. **Olaimat AN, Holley RA.** 2012. Factors influencing the microbial safety of fresh produce: A review. Food Microbiol. **32:**1-19.
6. **Warriner K, Huber A, Namvar A, Fan W, Dunfield K.** 2009. Recent Advances in the Microbial Safety of Fresh Fruits and Vegetables. Adv. Food Nutr. Res., Vol 57 **57:**155-208.
7. **Bidol SA, Daly ER, Rickert RE, Hill TA, Al Khaldi S, Taylor TH, Jr., Lynch MF, Painter JA, Braden CR, Yu PA, Demma L, Behravesh CB, Olson CK, Greene SK, Schmitz AM, Blaney DD, Gershman M.** 2007. Multistate outbreaks of *Salmonella* infections associated with raw tomatoes eaten in restaurants – United States, 2005-2006. MMWR **56:**909-911.
8. **Mody RK, Greene SA, Gaul L, Sever A, Pichette S, Zambrana I, Dang T, Gass A,**

Wood R, Herman K, Cantwell LB, Falkenhorst G, Wannemuehler K, Hoekstra RM, McCullum I, Cone A, Franklin L, Austin J, Delea K, Behravesh CB, Sodha SV, Yee JC, Emanuel B, Al-Khaldi SF, Jefferson V, Williams IT, Griffin PM, Swerdlow DL. 2011. National Outbreak of *Salmonella* Serotype Saintpaul Infections: Importance of Texas Restaurant Investigations in Implicating Jalapeno Peppers. Plos One **6:**7.

9. **US Food and Drug Administration.** 2012. Final Update on Multistate Outbreak of Listeriosis Linked to Whole Cantaloupes. http://www.fda.gov/food/foodsafety/corenetwork/ucm272372.htm

10. **US Food and Drug Administration.** 2011-2013, posting date. All Recalls. FDA. http://www.fda.gov/Safety/Recalls/default.htm

11. **Berger CN, Sodha SV, Shaw RK, Griffin PM, Pink D, Hand P, Frankel G.** 2010. Fresh fruit and vegetables as vehicles for the transmission of human pathogens. Environ Microbiol **12:**2385-2397.

12. **Beuchat LR.** 2006. Vectors and conditions for preharvest contamination of fruits and vegetables with pathogens capable of causing enteric diseases. Brit. Food J. **108:**38-53.

13. **Harris LJ, Farber JN, Beuchat LR, Parish ME, Suslow TV, Garrett EH, Busta FF.** 2003. Outbreaks Associated with Fresh Produce: Incidence, Growth, and Survival of Pathogens in Fresh and Fresh Cut Produce. Comp. Rev. Food Sci. Food Safety **2:**78-141.

14. **Strawn LK, Fortes ED, Bihn EA, Nightingale KK, Grohn YT, Worobo RW, Wiedmann M, Bergholz PW.** 2013. Landscape and meteorological factors affecting prevalence of three food-borne pathogens in fruit and vegetable farms. Appl. Environ. Microbiol. **79:**588-600.

15. **Beretti M, Stuart D.** 2008. Food safety and environmental quality impose conflicting demands on Central Coast growers. Calif. Agr. **62:**68-73.

16. **Lowell K, Langholz J, Stuart D.** 2010. Safe and Sustainable: Co-Managing for Food Safety and Ecological Health in California's Central Coast Region. The Nature Conservancy of California and the Georgetown University, Produce Safety Project, San Francisco, CA and Washington, D.C.

17. **Gennet S, Howard J, Langholz J, Andrews K, Reynolds MD, Morrison SA.** 2013. Farm practices for food safety: an emerging threat to floodplain and riparian ecosystems. Front. Ecol. Environ.**11**:236-242.

18. **Hardesty SD, Kusunose Y.** 2009. Growers' Compliance Costs for the Leafy Greens Marketing Agreement and Other Food Safety Programs. UC Small Farm Program Research Brief. ANR Univ. Calif.. http://sfp.ucdavis.edu/files/143911.pdf.

19. **Beuchat LR.** 1996. Pathogenic microorganisms associated with fresh produce. J. Food Prot. **59:**204-216.

20. **Cooley M, Carychao D, Crawford-Miksza L, Jay MT, Myers C, Rose C, Keys C, Farrar J, Mandrell RE.** 2007. Incidence and Tracking of *Escherichia coli* O157:H7 in a Major Produce Production Region in California. Plos One **2:**16.

21. **Jay MT, Cooley M, Carychao D, Wiscomb GW, Sweitzer RA, Crawford-Miksza L, Farrar JA, Lau DK, O'Connell J, Millington A, Asmundson RV, Atwill ER, Mandrell RE.** 2007. *Escherichia coli* O157 : H7 in feral swine near spinach fields and cattle, central California coast. Emerg. Infect. Dis. **13:**1908-1911.

22. **Foltz VD.** 1969. *Salmonella* Ecology. J. Am. Oil Chem. Soc. **46:**222-&.

23. **Gorski L, Parker CT, Liang A, Cooley MB, Jay-Russell MT, Gordus AG, Atwill ER, Mandrell RE.** 2011. Prevalence, Distribution, and Diversity of *Salmonella enterica* in a Major Produce Region of California. Appl. Environ. Microbiol. **77:**2734-2748.

24. **Haley BJ, Cole DJ, Lipp EK.** 2009. Distribution, Diversity, and Seasonality of

Waterborne *Salmonellae* in a Rural Watershed. Appl. Environ. Microbiol. **75:**1248-1255.

25. **Winfield MD, Groisman EA.** 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl. Environ. Microbiol. **69:**3687-3694.

26. **Jokinen C, Edge TA, Ho S, Koning W, Laing C, Mauro W, Medeiros D, Miller J, Robertson W, Taboada E, Thomas JE, Topp E, Ziebell K, Gannon VPJ.** 2011. Molecular subtypes of *Campylobacter* spp., *Salmonella enterica*, and *Escherichia coli* O157:H7 isolated from faecal and surface water samples in the Oldman River watershed, Alberta, Canada. Water Res. **45:**1247-1257.

27. **Patchanee P, Molla B, White N, Line DE, Gebreyes WA.** 2010. Tracking *Salmonella* Contamination in Various Watersheds and Phenotypic and Genotypic Diversity. Foodborne Pathog. Dis. **7:**1113-1120.

28. **Johnson JYM, Thomas JE, Graham TA, Townshend I, Byrne J, Selinger LB, Gannon VPJ.** 2003. Prevalence of *Escherichia coli* O157 : H7 and *Salmonella* spp. in surface waters of southern Alberta and its relation to manure sources. Can. J. Microbiol. **49:**326-335.

29. **Wilkes G, Edge TA, Gannon VPJ, Jokinen C, Lyautey E, Neumann NF, Ruecker N, Scott A, Sunohara M, Topp E, Lapen DR.** 2011. Associations among pathogenic bacteria, parasites, and environmental and land use factors in multiple mixed-use watersheds. Water Res. **45:**5807-5825.

30. **Manas P, Castro E, de las Heras J.** 2009. Irrigation with treated wastewater: Effects on soil, lettuce (*Lactuca sativa*) crop and dynamics of microorganisms. J. Environ. Sci. Health Part a-Toxic/Hazardous Subs. & Environ. Eng. **44:**1261-1273.

31. **Nightingale KK, Fortes ED, Nightingale CR, Her Z, Schukken YH, Grohn YT, Wiedmann M.** 2003. Transmission and ecology of *Listeria monocytogenes* in ruminant

hosts and the pre-harvest food environment. Abstr. Gen. Meet. Am. Soc. Microbiol. **103:**P-079.

32. **Lyautey E, Hartmann A, Pagotto F, Tyler K, Lapen DR, Wilkes G, Piveteau P, Rieu A, Robertson WJ, Medeiros DT, Edge TA, Gannon V, Topp E.** 2007. Characteristics and frequency of detection of fecal *Listeria monocytogenes* shed by livestock, wildlife, and humans. Can. J. Microbiol. **53:**1158-1167.

33. **Centers for Disease Control and Prevention.** 1999. Reptile-associated Salmonellosis–selected states, 1996-1998. MMWR. **48:**1009-1013.

34. **Cummings KJ, Warnick LD, Elton M, Rodriguez-Rivera LD, Siler JD, Wright EM, Groehn YT, Wiedmann M.** 2010. *Salmonella enterica* Serotype Cerro Among Dairy Cattle in New York: An Emerging Pathogen? Foodborne Pathog. Dis. **7:**659-665.

35. **Hoelzer K, Switt AIM, Wiedmann M.** 2011. Animal contact as a source of human non-typhoidal salmonellosis. Vet. Res. **42:**28.

36. **Centers for Disease Control and Prevention.** 2010. Multistate outbreak of human *Salmonella* typhimurium infections associated with aquatic frogs – United States, 2009. MMWR. **58:**1433-1436.

37. **Islam M, Morgan J, Doyle MP, Phatak SC, Millner P, Jiang XP.** 2004. Fate of *Salmonella enterica* serovar Typhimurium on carrots and radishes grown in fields treated with contaminated manure composts or irrigation water. Appl. Environ. Microbiol. **70:**2497-2502.

38. **Mukhejee A, Speh D, Diez-Gonzalez F.** 2007. Association of farm management practices with risk of *Escherichia coli* contamination in pre-harvest produce grown in Minnesota and Wisconsin. Int. J. Food Microbiol. **120:**296-302.

39. **Park S, Navratil S, Gregory A, Bauer A, Srinath I, Jun M, Szonyi B, Nightingale K,**

Anciso J, Ivanek R. 2013. Generic *Escherichia coli* Contamination of Spinach at the Preharvest Stage: Effects of Farm Management and Environmental Factors. Appl. Environ. Microbiol. **79:**4347-4358.

40. **Doyle MP, Erickson MC.** 2012. Opportunities for mitigating pathogen contamination during on-farm food production. Int. J. Food Microbiol. **152:**54-74.

41. **Park S, Szonyi B, Gautam R, Nightingale K, Anciso J, Ivanek R.** 2012. Risk Factors for Microbial Contamination in Fruits and Vegetables at the Preharvest Level: A Systematic Review. J. Food Prot. **75:**2055-2081.

42. **Uesugi AR, Danyluk MD, Mandrell RE, Harris LJ.** 2007. Isolation of *Salmonella* Enteritidis phage type 30 from a single almond orchard over a 5-year period. J. Food Prot. **70:**1784-1789.

43. **Enivironmental Protection Agency.** 2000. Improved Enumeration Methods for the Recreational Water Quality Indicators: *Enterococci* and *Escherichia coli*, vol. EPA/821/R-97/004. Office or Science and Technology, Washington, D.C.

44. **Environmental Protection Agency.** 2002. Method 1603: *Escherichia coli* (*E. coli*) in Water by Membrane Filtration Using Modified membrane-Thermotolerant *Escherichia coli* Agar (Modified mTEC), vol. EPA 821-R-02-023. Office of Science and Technology, Washington D.C.

45. **Andrews WH, Hammack T** 2011, posting date. Bacteriological Analytical Manual, Chapter 5, *Salmonella*. Food and Drug Administration. http://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm070149.htm.

46. **Hitchins AD, Jinneman K.** 2011. BAM: Detection and Enumeration of *Listeria monocytogenes*, April 2011 ed, vol. Ch. 10. Food and Drug Administration. http://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm071400.htm.

47. **Kim JS, Lee GG, Park JS, Jung YH, Kwak HS, Kim SB, Nam YS, Kwon ST.** 2007. A novel multiplex PCR assay for rapid and simultaneous detection of five pathogenic bacteria: *Escherichia coli* O157 : H7, *Salmonella*, *Staphylococcus aureus*, *Listeria monocytogenes*, and *Vibrio parahaemolyticus*. J. Food Prot. **70:**1656-1662.

48. **Nightingale KK, Windham K, Wiedmann M.** 2005. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. J. Bacteriol. **187:**5537-5551.

49. **den Bakker HC, Bundrant BN, Fortes ED, Orsi RH, Wiedmann M.** 2010. A population genetics-based and phylogenetic approach to understanding the evolution of virulence in the genus *Listeria*. Appl. Environ. Microbiol. **76:**6085-6100.

50. **Bundrant BN, Hutchins T, den Bakker HC, Fortes E, Wiedmann M.** 2011. Listeriosis outbreak in dairy cattle caused by an unusual *Listeria monocytogenes* serotype 4b strain. J. Vet. Diagn. Invest. **23:**155-158.

51. **Noah CW, Shaw CI, Ikeda JS, Kreuzer KS, Sofos JN.** 2005. Development of green fluorescent protein-expressing bacterial strains and evaluation for potential use as positive controls in sample analyses. J. Food Prot. **68:**680-686.

52. **Roberts AJ, Wiedmann M.** 2006. Allelic exchange and site-directed mutagenesis probe the contribution of ActA amino-acid variability to phosphorylation and virulence-associated phenotypes among *Listeria monocytogenes* strains. Fems. Microbiol. Lett. **254:**300-307.

53. **Grimont PAD, Weill FX.** 2007. Antigenic Formulae of the *Salmonella* Serovars, 9th ed. World Health Organization Collaborating Centre for Reference and Research on *Salmonella*, Pasteur Institute, Paris.

54. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.**

1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389-3402.

55. **Nightingale K, Bovell L, Grajczyk A, Wiedmann M.** 2007. Combined *sigB* allelic typing and multiplex PCR provide improved discriminatory power and reliability for *Listeria monocytogenes* molecular serotyping. J. Microbiol. Methods **68:**52-59.

56. **Hutchison ML, Walters LD, Moore A, Crookes KM, Avery SM.** 2004. Effect of length of time before incorporation on survival of pathogenic bacteria present in livestock wastes applied to agricultural soil. Appl. Environ. Microbiol. **70:**5111-5118.

57. **Solomon EB, Yaron S, Matthews KR.** 2002. Transmission of *Escherichia coli* O157 : H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. Appl. Environ. Microbiol. **68:**397-400.

58. **Holley RA, Arrus KM, Ominski KH, Tenuta M, Blank G.** 2006. *Salmonella* survival in manure-treated soils during simulated seasonal temperature exposure. J. Environ. Qual. **35:**1170-1180.

59. **Ohtomo R, Minato K, Saito M.** 2004. Survival of *Escherichia coli* in a field amended with cow feces slurry. Soil Sci. Plant Nutr. **50:**575-581.

60. **You Y, Rankin SC, Aceto HW, Benson CE, Toth JD, Dou Z.** 2006. Survival of *Salmonella enterica* serovar Newport in manure and manure-amended soils. Appl. Environ. Microbiol. **72:**5777-5783.

61. **Franz E, Semenov AV, van Bruggen AHC.** 2008. Modelling the contamination of lettuce with *Escherichia coli* O157:H7 from manure-amended soil and the effect of intervention strategies. J. Appl. Microbiol. **105:**1569-1584.

62. **Franz E, Semenov AV, Termorshuizen AJ, de Vos OJ, Bokhorst JG, van Bruggen AHC.** 2008. Manure-amended soil characteristics affecting the survival of *E-coli* O157 :

H7 in 36 Dutch soils. Environ. Microbiol. **10:**313-327.

63. **Semenov AV, van Overbeek L, van Bruggen AHC.** 2009. Percolation and Survival of *Escherichia coli* O157:H7 and *Salmonella enterica* Serovar Typhimurium in Soil Amended with Contaminated Dairy Manure or Slurry. Appl. Environ. Microbiol. **75:**3206-3215.

64. **Natvig EE, Ingham SC, Ingham BH, Cooperband LR, Roper TR.** 2002. *Salmonella enterica* serovar Typhimurium and *Escherichia coli* contamination of root and leaf vegetables grown in soils with incorporated bovine manure. Appl. Environ. Microbiol. **68:**2737-2744.

65. **Lung AJ, Lin CM, Kim JM, Marshall MR, Nordstedt R, Thompson NP, Wei CI.** 2001. Destruction of *Escherichia coli* O157 : H7 and *Salmonella Enteritidis* in cow manure composting. J. Food Prot. **64:**1309-1314.

66. **Leafy Green Marketing Agreenment Advisory Board.** 2012. Commodity Specific Food Safety Guidelines for the Production and Harvest of Lettuce and Leafy Greens. http://www.lgma.ca.gov/sites/default/files/08.31.12%20CALGMA%20GAPs%20-%20metrics.pdf.

67. **Koelsch RK, Lorimor JC, Mankin KR.** 2006. Vegetative treatment systems for management of open lot runoff: Rev. Lit. Appl. Eng. Agr. **22:**141-153.

68. **Hill VR, Sobsey MD.** 2001. Removal of *Salmonella* and microbial indicators in constructed wetlands treating swine wastewater. Water Sci. Technol. **44:**215-222.

69. **Stuart D, Shennan C, Brown M** 2006. Food Safety versus Environmental Protection on the Central California Coast: Exploring the Science Behind an Apparent Conflict Center for Agroecol. Sustain. Food Sys. http://casfs.ucsc.edu/wp-content/uploads/2010/10/Research_Brief_10.pdf.

70. **Langholz J, Jay-Russell M.** 2013. Potential role of wildlife in pathogenic contamination of fresh produce. Hum. Wildl. Interact. **7:**140–157.

71. **Sivapalasingam S, Friedman CR, Cohen L, Tauxe RV.** 2004. Fresh produce: A growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. J. Food Prot. **67:**2342-2353.

72. **Steele M, Odumeru J.** 2004. Irrigation water as source of foodborne pathogens on fruit and vegetables. J. Food Prot. **67:**2839-2849.

73. **Fonseca JM, Fallon SD, Sanchez CA, Nolte KD.** 2011. *Escherichia coli* survival in lettuce fields following its introduction through different irrigation systems. J. Appl. Microbiol. **110:**893-902.

74. **Gu GY, Luo ZY, Cevallos-Cevallos JM, Adams P, Vellidis G, Wright A, van Bruggen AHC.** 2013. Factors affecting the occurrence of *Escherichia coli* O157 contamination in irrigation ponds on produce farms in the Suwannee River Watershed. Can. J. Microbiol. **59:**175-182.

75. **Benjamin L, Atwill ER, Jay-Russell M, Cooley M, Carychao D, Gorski L, Mandrell RE.** 2013. Occurrence of generic *Escherichia coli*, *E. coli* O157 and *Salmonella* spp. in water and sediment from leafy green produce farms and streams on the Central California coast. Int. J. Food Microbiol. **165:**65-76.

76. **Lyautey E, Lapen DR, Wilkes G, McCleary K, Pagotto F, Tyler K, Hartmann A, Piveteau P, Rieu A, Robertson WJ, Medeiros DT, Edge TA, Gannon V, Topp E.** 2007. Distribution and characteristics of *Listeria monocytogenes* isolates from surface waters of the South Nation River watershed, Ontario, Canada. Appl. Environ. Microbiol. **73:**5401-5410.

77. **Sauders B.** 2005. Molecular epidemiology, diversity, distribution and ecology

of *Listeria*. PhD. Cornell University, Ithaca, NY.

78. **Rudolfs W, Falk LL, Ragotzkie RA.** 1951. Contamination of Vegetables Grown in Polluted Soil.1. Bacterial Contamination. Sewage Ind. Wastes **23:**253-268.

79. **Welshimer HJ.** 1960. Survival of *Listeria monocytogenes* in Soil. J. Bacteriol. **80:**316-320.

CHAPTER 4

DISTRIBUTION OF *SALMONELLA* SUBTYPES DIFFERS BETWEEN TWO US PRODUCE GROWING REGIONS

ABSTRACT

Salmonella accounts for approximately 50% of the produce associated outbreaks in the United States (US), many of which traceback to contamination in the production environment. To aid in identification of these contamination sources, we characterized *Salmonella* isolates from two geographical diverse produce growing regions in the US. Initially, we characterized the *Salmonella* serotype and subtypes diversity associated with 1,677 samples collected from 33 produce farms New York State (NYS). Among these 1,677 samples, 74 were *Salmonella* positive, yielding 80 unique isolates that represented 14 serovars and 23 different PFGE-types. To explore regional *Salmonella* diversity associated with production environments, we collected a smaller set of samples (n=65) from South Florida (S. FL.) production environments and compared the *Salmonella* diversity associated with these samples with the diversity found among NYS production environments. Among these 65 samples, 23 were *Salmonella* positive, yielding 32 unique isolates that represented 11 serovars and 17 different PFGE-types. The most common serovars isolated in NYS were Newport, Cerro and Thompson, while common serovars isolated in S. FL were Saphra and Newport, and *S. enterica* subsp. *diarizonae* serovar 50:r:z. A high PFGE-type diversity (D=0.90±0.02) was observed amongst *Salmonella* isolates across both regions; only three PFGE-types were shared among the two regions. The probability of three or fewer shared PFGE-types was <0.000001; therefore, *Salmonella* isolates were considerably different between the two sampled regions. These findings suggest the potential for PFGE-based source tracking of *Salmonella* in production environments.

INTRODUCTION

The genus *Salmonella* is divided into two species, *S. enterica* and *S. bongori* and represents approximately 2,600 known serovars. *S. enterica* has six subspecies: *enterica* (I) *salamae* (II), *arizonae* (IIIa), *diarizonae* (IIIb), *houtanae* (IV), and *indica* (V) and accounts for over 99% of *Salmonella* strains isolated worldwide (1). *S. enterica* subsp. *enterica* is the leading cause of bacterial foodborne illnesses, hospitalizations, and deaths in the United States (US) (2). Approximately 95% of *S. enterica* infections in the US are a result of consumption of contaminated foods (3, 4). Furthermore, *S. enterica* is estimated to be responsible for half of the produce associated illnesses and a majority of produce associated outbreaks in the US (4-7). As a result, there is a need to characterize *Salmonella* isolated throughout the farm to fork continuum to help identify likely sources of produce contamination.

Subtyping is an important tool to detect foodborne outbreaks and to identify outbreak sources (8), as well as a powerful tool to investigate the diversity of foodborne pathogens in various hosts and environments. Even the most basic of subtyping methods (e.g., serotyping) can yield information on likely reservoirs for specific foodborne pathogens (9, 10). For instance, *S.* Dublin and Choleraesuis are routinely associated with cattle and swine hosts, respectively (10, 11). In other studies (1, 12), *S. enterica* subsp. *salamae*, *arizonae*, *diarizonae*, *houtanae*, and *indica* were found to be predominantly associated with cold blooded animal hosts. For example, *S. enterica* subsp. *diarizonae* was isolated most frequently from reptiles and amphibians obtained from the central coast of California (US) (13). Another study (14) found approximately 81% of pet snake fecal samples collected in Germany were positive for *S. enterica* subsp. *diarizonae* serovars. Thus in cases where a single subtype of *Salmonella*, known to be host restricted or host adapted, is detected in an environment, its isolation in a food may indicate a specific point source of contamination. Contamination may be linked to a single/small subset of hosts that carry the

subtype. However, some *S. enterica* serovars are associated with a broad host range and a number of geographically diverse regions. Serovars Typhimurium and Enteritidis are two of the most common serovars reported in human *Salmonella* isolates worldwide (15). Though, a study conducted in Great Britain on wild bird populations was able to identify two host-adapted *S.* Typhimurium strains by use of pulsed field gel electrophoresis (PFGE) and phage typing (PT) (16), demonstrating the application of subtyping and host associations even in broad host range serovars.

While subtyping can yield likely information on potential hosts of *Salmonella*, studies (17-21) have also demonstrated the application of subtyping to track specific foodborne pathogen subtypes in the environment. For instance, one study (21) investigating the source of fecal pollution in a Japanese beach used PFGE-typing to show a strong association between *Enterococcus faecium* isolated in samples from the beach and one of the suspected contamination sources (a river that drains into the beach). PFGE-typing was also used to track *E. coli* O157:H7 in a California produce growing region. The same PFGE-type of *E. coli* O157:H7 was isolated from feral swine, cattle, surface water, sediment and soil from one of the spinach farms that had been implicated in the 2006 spinach-borne *E. coli* O157:H7 outbreak. In another study, Patchanee *et al.* (19) reported that distinct *Salmonella* PFGE-types were recovered from water samples collected from a site near swine production or forestry, residential/industrial, and agriculture cropland, further supporting that subtyping methods, such as PFGE-typing, can be used to track *Salmonella* in the environment and to identify specific contamination sources.

Studies (19, 22-24) have characterized the distribution and diversity of *Salmonella* from a number of different environments; however, there is minimal information on *Salmonella* in the produce production environment, and no one to our knowledge has compared *Salmonella* isolated from two geographically diverse produce growing regions using the same sample

collection, detection, and isolation schemes. The purpose of this study was to characterize *Salmonella* isolates obtained from environmental samples collected in produce production environments in New York State (NYS) and South Florida (S. FL). Specifically, we used both qualitative and quantitative methods to examine the distribution and diversity of *Salmonella* isolates from each region, as well as to compare the subtype distribution between these two regions. Additionally, *Salmonella* isolate subtype data was used to suggest potential sources of *Salmonella* contamination.

MATERIALS AND METHODS

Description of isolates used in this study. A total of 228 *Salmonella* isolates (147 isolates from NYS and 81 isolates from S. FL) were assembled for this study using five *Salmonella* datasets, two published, two unpublished, and one reported here (Table 1). Four datasets (I, II, IV, and V) representing NYS production environments had been collected to explore the association between *Salmonella* prevalence in produce fields and geographical and/or management factors, whereas one dataset (III) obtained from S. FL was specifically collected for this study. All sample collection and preparation (i.e., preparation of samples for *Salmonella* enrichment) was performed using the same methodology as previously described (25, 26). Detection and isolation of *Salmonella* was performed using a modified version of the Food and Drug Administration Bacteriological Analytic Manual (FDA BAM) (27). Briefly, samples were diluted 1:10 with tryptic soy broth (TSB; Becton Dickinson; Franklin Lakes, NJ) and incubated for 24 h at 35±2°C. Enrichment aliquots of 1.0 and 0.1 mL were transferred to tetrathionate (TT; Oxoid; Cambridge, United Kingdom) and Rappaport Vassiliadis (RV; Oxoid) broths, respectively, and incubated for 24 h in a shaking water bath at 42±2°C. TT and RV were plated onto xylose lysine deoxycholate agar (XLD; Neogen; Lansing, MI) and *Salmonella* chromogenic agar

(CHROMagar; CHROMagar Company; Paris, France), which were incubated for 24 and 48 h at 35 and 37±2°C, respectively. Presumptive *Salmonella* colonies (up to four colonies per isolation scheme, e.g., TT-XLD, RV-XLD, TT-CHROMagar, and RV-CHROMagar) were sub-streaked to brain heart infusion (BHI; Becton Dickinson) agar and incubated for 24 h at 37±2°C. One isolated colony was selected from each BHI plate and confirmed by polymerase chain reaction (PCR) for *invA*, which is specific to *Salmonella* (28). All confirmed *Salmonella* isolates were preserved at -80°C in 15% glycerol.

Table 4.1 Summary of Study Datasets and *Salmonella* Isolates[a]

Dataset	Region[b]	Year(s)	Sample Type	No. Samples	Frequency of positive samples (%)	Isolates Subtyped[c]	No. Rep. Isolates[d]	Serovars	PFGE types
I	NYS	2009-2011	Total	**588**	**27 (4.6)**	57[e]	27	7	11
			Soil	178	4 (2.8)				
			Drag swab	175	3 (1.7)				
			Water	174	16 (9.2)				
			Fecal	61	4 (6.6)				
II	NYS	2012	Total	**600**	**26 (4.3)**	35	29	9	11
			Soil	263	13 (4.9)				
			Drag swab	263	5 (1.9)				
			Water	74	8 (10.8)				
			Fecal	NC[f]					
III	S. FL	2010	Total	**65**	**23 (35.4)**	81	32	11	17
			Soil	8	5 (62.5)				
			Drag swab	8	3 (37.5)				
			Water	40	15(37.5)				
			Fecal	9	0 (0)				
IV	NYS	2011	Total	**429**	**17 (4.0)**	44	19	7	9
			Soil	90	2 soil (2.2)				
			Drag swab	219	4 drag swab (1.8)				
			Water	120	11 water (9.2)				
			Fecal	NC					
V	NYS	2010	Total	**60**	**5 (8.3)**	11	5	4	4
			Soil	20	2 (10)				
			Drag swab	20	0 (0)				
			Water	13	1 (7.7)				
			Fecal	7	2 (28.6)				
All	–	–	Total	**1,742**	**98 (5.8)**	228	112	20	37
			Soil	559	26 (5.1)				
			Drag swab	685	15 (2.2)				
			Water	421	51 (12.1)				
			Fecal	77	6 (7.8)				

[a] Dataset I *Salmonella* isolates, serovars, and PFGE types have been previously described (25); Dataset II *Salmonella* isolates and serovars have been previously described (26); Dataset III is reported here; Dataset IV and V *Salmonella* isolates and serovars are

unpublished. See supplemental material (S1) for further dataset descriptions.
[b] Region abbreviations NYS and S. FL stand for New York State and S. FL, respectively.
[c] PFGE (using the restriction enzyme *Xba*I) was performed on one isolate per *Salmonella* positive sample for each isolation scheme (up to four isolates may be selected (RV-XLD, TT-XLD, RV-CHROMagar, and TT-CHROMagar)).
[d] Only isolates that were representative(s) of the *Salmonella* isolated in that sample were kept for further analyses (i.e., isolates from the same sample with identical serovar and PFGE type were excluded).
[e] Isolates from one *Salmonella* positive sample were unavailable for subtyping.
[f] NC represents the sample type was not collected for that dataset.
.

Isolates selected for molecular characterization. One isolate per isolation scheme (e.g., TT-XLD, RV-XLD, TT-CHROMagar, and RV-CHROMagar) was selected for serotyping and PFGE-typing (147 isolates from NYS and 81 isolates from S. FL; n=228 isolates). This approach was used to capture all potential strains of *Salmonella* that may be present in a sample, as several studies (13, 16, 22, 26, 29, 30) have shown that multiple *Salmonella* strains may be isolated from the same sample.

Serotyping. Traditional serotyping was performed on all 288 isolates by the Wadsworth Center, New York State Department of Health (Albany, NY) using the White-Kauffman-Le Minor scheme (1).

Pulsed Field Gel Electrophoresis. PFGE-typing was performed on all 228 isolates using the standard Centers for Disease Control and Prevention (CDC) PulseNet protocol (31). Briefly, *Salmonella* cells were embedded in 1% SeaKem Gold agarose (SeaKem Gold Agarose; Lonza; Rockland, ME), lysed, washed, and digested with 50 U/plug of *Xba*I (Roche Applied Science; Indianapolis, IN) at 37°C. Separation of the restricted DNA fragments was performed by a Chef Mapper® XA (Bio-Rad; Hercules, California) for 18-20 h in 1% agarose gels. Voltage was set to 6 V/cm with an initial switch time of 2.16 s and final switch time of 63.8 s. *S. enterica* subsp. *enterica* Braenderup was used as the reference standard to allow for normalization and comparison of gel images (32). Gel images were captured by a Bio-Rad Gel Doc using Multi-Analyst software 1.1 (Bio-Rad). PFGE images were analyzed by BioNumerics software 5.1 (Applied Maths; Sint-Martens-Latem, Belgium). Similarity clustering analyses were performed using the unweighted pair group method with arithmetic mean algorithm (UPGMA) based on Dice coefficients with a maximum space tolerance of 1.5%. PFGE-types were named with the standard CDC PulseNet nomenclature and were assigned by comparison of patterns to an internal reference database (Food Safety Laboratory (FSL) Cornell University, Ithaca, NY)

comprised of approximately 6,000 *Salmonella* isolates. Lastly, PFGE was used to predict serovar by comparison to the FSL database using a similarity cluster analysis (as described above).

Molecular serotyping. Molecular serotyping was performed on 12 isolates where serovars reported by traditional serotyping did not match serovars predicted by PFGE. Serovar was confirmed by PCR detection of the *Salmonella* O serogroup genes by a multiplex PCR assay that simultaneously targets genes for five *Salmonella* O antigens (B, C1, C2-C3, D1, and E1), as described by Ranieri *et al.* (9) and Herrera-Léon *et al.* (33); in addition, PCR amplification and sequencing of the genes encoding the H1 and H2 antigens was performed. Primer sets, DNA amplification, and PCR conditions for *fliC* (encodes H1 antigen) and *fljB* (encodes H2 antigen) have been previously described by Imre *et al.* (34, 35) and Mortimer *et al.* (34, 35). DNA sequence data was compared to sequences in an internal database (FSL) of H1 and H2 sequences, as previously described (9). A serovar was identified for each of the 12 isolates; therefore, molecular serotyping results were able to resolve serovar conflicts for the isolates with different serovars identified by traditional serotyping and PFGE.

Final *Salmonella* isolate dataset. Upon completion of serotyping and PFGE-typing, only isolates that were "unique" (i.e., representative(s) of the *Salmonella* isolated in that sample) were kept for further analyses; isolates from the same sample with identical serovar and PFGE-type were excluded from further analysis. This approach yielded 112 representative *Salmonella* isolates (80 isolates from NYS and 32 isolates from S. FL).

Statistical Analysis. All analyses were performed in the statistical computing environment, R 3.0.2 (36). Diversity of PFGE-type between the NYS and S. FL regions was assessed, using two assessment tools: a diversity index and probability simulations.

Simpson's Index of Diversity (D) was calculated amongst all PFGE-types, and for each region (37), with 95% confidence intervals (38). D values closest to 1 indicate a high diversity,

while D values closest to 0 indicate a low diversity. To test if isolates from NYS and S. FL were drawn from two distinct populations, we performed a series of simulations to quantify the likelihood of three or fewer shared subtypes occurring between the regions, given that all the *Salmonella* isolates were drawn from one population. Briefly, we randomly permuted the isolates across the two regions and computed the number of shared PFGE types. The simulation was performed 1,000,000 times and the probability of simulations (i.e., permutations) with three or fewer shared subtypes was calculated.

Data Access. Isolate information and subtyping data from this study are archived and available through the Food Microbe Tracker database (http://www.foodmicrobetracker.com).

RESULTS AND DISCUSSION

Our study is one of the first to characterize the *Salmonella* distribution and diversity associated with the produce production environment in NYS. In addition, we assembled a smaller set of *Salmonella* isolates from S. FL that were characterized to examine the regional distribution of subtypes between NYS and S. FL. In the NYS dataset, 74 of 1,677 previously collected environmental samples were *Salmonella* positive (Table 1). The 74 *Salmonella* positive samples yielded 147 isolates. Upon completion of serotyping and PFGE-typing, 80 isolates that were representative(s) of the *Salmonella* isolated in that sample were retained for further analyses and discussion. In the smaller S. FL dataset, 23 of 65 previously collected environmental samples were *Salmonella* positive (Table 1). The 23 *Salmonella* positive samples yielded 81 isolates, which were condensed to 32 representative isolates. The combined NYS and S. FL datasets consisted of 112 representative *Salmonella* isolates (Table 2). Overall, multiple serovars from *Salmonella* positive samples were isolated across both NYS and S. FL datasets. Additionally, across both NYS and S. FL datasets, we observed a number of different *S. enterica*

serovars that were associated with certain hosts. Lastly, we showed a distinct difference between *Salmonella* isolates from the two produce growing regions (NYS and S. FL), as determined by analysis of PFGE types.

Table 4.2. Serovars and PFGE types found among *S. enterica* isolates from study samples collected in New York State (NYS) and S. FL (S. FL) US produce production environments.

		NYS datasets					S. FL dataset
Serovar[a]	PFGE type[b]	I	II	IV	V	Total	III (Total)
S. enterica subsp. *enterica*							
Agona	NYCU.JAAX01.1131	–[c]	3	–	–	3	–
Baildon	NYCU.JAAX01.0345	–	–	–	–	–	2
Braenderup	NYCU.JAAX01.1196	–	–	–	–	–	1
Cerro	NYCU.JAAX01.0213	10	5	–	1	16	–
Enteritidis	NYCU.JAAX01.1225	–	–	4	–	4	–
Gaminara	NYCU.JAAX01.1202	–	–	–	–	–	1
Give	NYCU.JAAX01.1215	1	–	–	–	1	–
	NYCU.JAAX01.1216	1	2	–	–	3	–
	NYCU.JAAX01.1217	–	–	1	–	1	–
	NYCU.JAAX01.1218	–	–	1	–	1	–
Infantis	NYCU.JAA.X01.1203	–	–	2	–	2	–
Litchfield	NYCU.JAAX01.1197	–	–	–	–	–	1
	NYCU.JAAX01.1198	–	–	–	–	–	1
Newport	NYCU.JAAX01.0121	–	6	–	–	6	–
	NYCU.JAAX01.0126	–	1	–	–	1	–
	NYCU.JAAX01.0296	2	2	2	1	7	–
	NYCU.JAAX01.1212	1	–	4	–	5	–
	NYCU.JAAX01.1213	–	–	–	–	–	1
	NYCU.JAAX01.1221[d]	3	–	–	–	3	2
	NYCU.JAAX01.1222	–	–	–	–	–	3
	NYCU.JAAX01.1223	–	–	–	–	–	1
Rubislaw	NYCU.JAAX01.1201	–	–	–	–	–	3
	NYCU.JAAX01.1220	–	–	1	–	1	–
Saphra	NYCU.JAAX01.1194	–	–	–	–	–	7
Senftenberg	NYCU.JAAX01.1005	–	1	–	–	1	–
Tennessee	NYCU.JAAX01.1214	–	1	–	–	1	–
Thompson	**NYCU.JAAX01.0157**	4	5	1	2	12	1
	NYCU.JAAX01.1199	1	–	–	–	1	1
	NYCU.JAAX01.1200	–	–	–	–	–	1
Typhimurium	NYCU.JAAX01.0072	–	–	–	–	–	1

Table 4.2 Continuation							
	NYCU.JAAX01.1207	–	1	–	–	1	–
	NYCU.JAAX01.1208	1	–	–	–	1	–
4,5,12:i:-	NYCU.JAAX01.1209	–	–	–	–	–	2
6,8:i:-	NYCU.JAAX01.0096	1	–	–	–	1	–
S. enterica subsp. *diarizonae*							–
50:r:z	NYCU.JAAX01.1210	–	–	3	–	3	–
	NYCU.JAAX01.1211	–	–	–	–	–	3
S. enterica subsp. *houtanae*							
40:z4,z32:-							
	NYCU.JAAX01.1219	2	2	–	1	5	–

[a] Serotyping was performed by agglutination at the Wadsworth Center, NYS Department of Health.
[b] PFGE was performed in accordance with the standard CDC PulseNet protocol for *Salmonella* using the restriction enzyme, *Xba*I (31). PFGE types were named with the standard CDC PulseNet nomenclature and were assigned by comparison of PFGE patterns to an internal reference database (Food Safety Laboratory, Cornell University, Ithaca, NY) comprised of approximately 6,000 *Salmonella* isolates.
[c] Dash (–) represents zero isolates.
[c] Bolded PFGE types (NYCU.JAAX01.1221, *S. enterica* subsp. *enterica* Newport; and NYCU.JAAX01.0157 and NYCU.JAAX01.1199, *S. enterica* subsp. *enterica* Thompson) were found in both NYS and S. FL.

Multiple serovars were isolated from *Salmonella* positive samples. Serotyping and PFGE typing were performed on one isolate per isolation scheme from the 97 *Salmonella* positive samples (74 and 23 *Salmonella* positive samples from NYS and S. FL, respectively). Of the 97 *Salmonella* positive samples, 83 samples yielded one serovar and 14 samples (approximately 15% of positive samples) yielded two or more serovars. The 14 *Salmonella* positive samples that yielded two or more serovars represented 6 and 8 samples from NYS and S. FL regions, respectively (Table 3). In the NYS region, two serovars were isolated in each of the six *Salmonella* positive samples. In the S. FL region, two serovars were isolated in seven of the eight *Salmonella* positive samples. Three serovars (*S.* Newport, *S.* Saphra, *S. enterica* subsp. *diarizonae* serovar 50:r:z) were isolated from the remaining one *Salmonella* positive sample (Table 3).

Several studies (13, 16, 22, 39) have isolated multiple *Salmonella* serovars from one sample. For instance, Jokinen (22) isolated more than one serovar from all *Salmonella* positive water samples (n=29) collected from a Canadian watershed. In our study, 50% of the *Salmonella* positive samples with multiple serovars represented water samples (3/6 and 4/8 positive samples from NYS and S. FL, respectively). Our water sample volume tested was 250 mL, while in the Jokinen et al. study the water sample volume tested was 500 mL. The number of serovars isolated from each sample may be influenced by the sample volume tested; however, further studies are required to correlate the volume of water tested and likely number of serovars. Isolation of multiple serovars from water samples may also be more likely because *Salmonella* may be more uniformly dispersed in water, compared to other samples types in our study (e.g., soil samples; 25 g).

Multiple *Salmonella* enrichment/plating schemes were used to limit *Salmonella* with atypical phenotypic characteristics from going undetected, as previous studies (29, 30, 40-43)

have shown that detection of certain serovars of *Salmonella* may be influenced by enrichment/plating media. In our study, the 14 *Salmonella* positive samples, where multiple serovars of *Salmonella* were isolated, yielded 14 different serovars (Table 3). Overall, no apparent association was observed between isolation of a *Salmonella* serovar and a specific media scheme (Table 3). For example, S. *enterica* subsp. *diarizonae* 50:r:z was isolated from three of the four media schemes (RV-CHROMagar, RV-XLD, and TT-XLD). Similarly, *S.* Saphra was also isolated from three of the four media schemes (TT-CHROMagar, RV-CHROMagar, and TT-XLD). Other serovars, that were isolated more than once, were also isolated from two or more media schemes. Select strains of four of the serovars isolated here (*S.* Thompson, Enteritidis, Typhimurium, and subsp. *diarizonae*) have been reported to be weak H_2S producers. No evident association was observed between serovars that have been reported as weak H_2S producers and a specific media scheme (Table 3). In this study, we used one plating medium that does (XLD) and does not (CHROMagar) use the production of H_2S as an indicator of *Salmonella* (41, 44). Additionally, select strains of six of the serovars isolated here (*S.* Tennessee, Newport, Agona, Senftenberg, Typhimurium, and subsp. *diarizonae*) have been reported to ferment lactose (45-47). Similar to findings for serovars that have been reported to be weak H_2S producers, no apparent association was observed here between the serovars that have been reported to ferment lactose and a specific media scheme (Table 3). Even though our sample size was small (14 cases of isolation of multiple *Salmonella* serovars from one sample), our results further support that the use of different enrichment/plating media schemes are needed to detect and isolate different *Salmonella* subtypes that may be present in a samples. This is supported by the observation that in a number of cases different *Salmonella* strains or serovars were isolated from different enrichment and plating procedures.

Table 4.3. Two or more serovars that were isolated from *Salmonella* positive samples by the four different detection and isolation schemes[a].

Sample	Region[b]	Sample Type	RV		TT	
			CHROMagar	XLD	CHROMagar	XLD
1	NYS	Soil	Newport	-	Thompson	-
2	NYS	Soil	Senftenberg	Newport	-	-
3	NYS	Drag swab	Tennessee	Agona	-	-
4	NYS	Water	-	-	Typhimurium	Give
5	NYS	Water	-	Enteritidis	-	Subsp. *diarizonae*
6	NYS	Water	Newport	-	-	Subsp. *diarizonae*
7	S. FL	Soil	Braenderup	-	Litchfield	-
8	S. FL	Soil	-	Newport	-	Saphra
9	S. FL	Drag swab	-	-	Thompson	Litchfield
10	S. FL	Drag swab	Baildon	-	Thompson	-
11	S. FL	Water	Newport	-	Saphra	-
12	S. FL	Water	Subsp. *diarizonae*	Thompson	-	-
13	S. FL	Water	Rubislaw	-	-	Gaminara
14	S. FL	Water	Saphra	Subsp. *diarizonae*	-	Newport

[a] Enrichment media: RV and TT represent Rappaport Vassiliadis and Tetrathionate broths, respectively. Plating media: CHROMagar and XLD represent *Salmonella* Chromogenic and Xylose Lysine deoxycholate agars, respectively.
[b] NYS and S. FL represent New York State and South Florida, respectively.

***Salmonella* serovars isolated from produce growing regions were diverse.** Combined the NYS and S. FL datasets represented 20 different serovars among the 112 subtyped representative isolates (Table 2). Specifically, 14 and 11 serovars were identified among the NYS (n=80) and S. FL (n=32) isolates, and five serovars were shared between the two regions. In total, *S. enterica* subsp. identified in both NYS and S. FL datasets were *enterica* (18 serovars), *diarizonae* (1 serovar) and *houtanae* (1 serovar) (Table 2). While *S. enterica* subsp. *enterica* serovars are often associated with warm blooded animal hosts and are the most commonly associated with human *Salmonella* infections linked to foods (48-50); a few documented cases have linked cold blooded animal hosts (e.g., reptiles, amphibians) to *S. enterica* subsp. *enterica* infections (10, 51, 52). However, serovars of *S. enterica* subsp. *diarizonae* and *houtanae* are primarily associated with cold blooded animal hosts (1, 12, 13); in addition, they are most commonly associated with human *Salmonella* infections linked to reptiles (14, 53). In the NYS dataset, isolates identified as *S. enterica* subsp. *enterica* represented serovars Newport (22 isolates), Cerro (16 isolates), Thompson (13 isolates), Give (6 isolates), Enteritidis (4 isolates), Agona (3 isolates), Typhimurium (2 isolates), Infantis (2 isolates), as well as 6,8:i:-, Rubislaw, Senftenberg, and Tennessee (1 isolate each); and *S. enterica* subsp. *diarizonae* and *houtanae* serovars 50:r:z (3 isolates) and 40:z4,z32:- (5 isolates), respectively (Table 2). In the S. FL dataset, isolates identified as *S. enterica* subsp. *enterica* represented serovars Newport (7 isolates), Saphra (7 isolates), Rubislaw (3 isolates), Thompson (3 isolates), 4,5,12:i:- (2 isolates), Litchfield (2 isolates), Baildon (2 isolates), as well as Braenderup, Gaminara, and Typhimurium (1 isolate each); and *S. enterica* subsp. *diarizonae* serovar 50:r:z (3 isolates) (Table 2). Serovars Newport, Cerro and Thompson were isolated most frequently in NYS and represented approximately 28, 20, and 16% of the NYS isolates (Table 2). On the other hand, serovars Newport and Saphra

were isolated most frequently in S. FL and represented 22% each of the S. FL isolates (Table 2). All other serovars represented <10% of the NYS and S. FL isolates.

Among the *Salmonella* positive samples obtained from NYS and S. FL, *S.* Newport was isolated the most (Fig 1; 22/80 isolates NYS and 7/32 isolates S. FL). *S.* Newport has a broad host range (e.g., dairy cattle, snakes, hedgehogs) and has been isolated from a diverse number of environmental sources (e.g., soil, water) (19, 54). Additionally, *S.* Newport is one of the most frequently reported serovars among human *Salmonella* isolates in North America, Europe, and Latin America (15, 55) and is associated with a wide range of foods and animals (e.g., beef, poultry) (50, 56). Produce-borne outbreaks of *S.* Newport have been routinely traced back to the produce production environment (57, 58), and have been associated with numerous commodities, including lettuce, mangoes, melons, alfalfa sprouts, and tomatoes (4, 59). For example, in 2005, tomatoes from Virginia were implicated as the food vehicle for an outbreak of *S.* Newport, with an estimated 72 cases of illness across 16 states. The source of this outbreak was traced back to a pond that served as an irrigation source for the tomato fields (58). The identification of *S.* Newport in both regions highlights the widespread nature of this serovar; in addition, the number of outbreaks associated with different produce commodities emphasizes the importance of this serovar in produce production environments.

S. Cerro was isolated exclusively from NYS (Fig. 1; 16/80 isolates). Previous studies (60-62) have indicated *S.* Cerro is prevalent in Pennsylvania (US) and NYS, especially in dairy production environments. This suggests *S.* Cerro may be a likely contaminant in produce production environments near dairy cattle operations. However, *S.* Cerro has rarely been associated with human *Salmonella* infections, and has been isolated from asymptomatic, healthy children in India (63). Of the few documented *S.* Cerro outbreaks, the most notable occurred in

New Mexico (US; 29 illnesses and 7 hospitalizations) and was linked to consumption of contaminated Carne Seca (63).

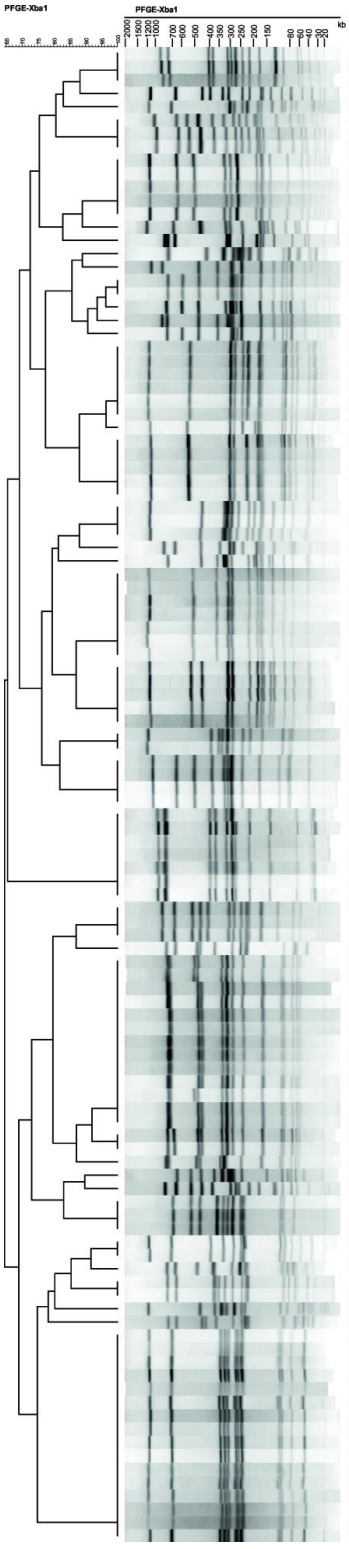
While serovar Thompson was isolated from both NYS and S. FL (Fig. 1; 13/80 isolates NYS and 3/32 isolates S. FL), it was more prevalent in NYS. One field study (64) conducted in NYS observed 10% (13/129) of surface water samples tested were positive for *S.* Thompson. This serovar has been associated with two produce-borne outbreaks; 1999 and 2004, there were 41 and 21 reported cases of salmonellosis linked to the consumption of contaminated cilantro in California, US (65) and rucola lettuce in Norway (66), respectively. While *S.* Thompson has only accounted for two known produce-borne outbreaks, it is common among human salmonellosis cases, as evident by its 13th rank in the CDC's top 20 *Salmonella* isolates from human sources (48).

S. Saphra was isolated exclusively from S. FL and represented 22% of the S. FL *Salmonella* isolates (Fig. 1; 7/32 isolates). This serovar is not commonly isolated from humans or animals, and thus minimal information has been reported on its ecology in the environment. Two studies, conducted in Argentina (67) and Brazil (68), isolated *S.* Saphra from surface water and animal drinking water samples. There has been only one documented outbreak of *S.* Saphra in the US, with 24 illnesses and 5 hospitalization, due to consumption of contaminated cantaloupe (69), purchased by a single distributor who obtained the cantaloupe from a specific region in Mexico.

Overall, we identified a number of *Salmonella* serovars that have previously been linked to produce-borne outbreaks; however, we also identified serovars that are rarely linked to human cases of salmonellosis (e.g., *S.* Cerro). We also observed several serovars that are associated with specific hosts, such as serovar 50:r:z (subsp. *diarizonae*), which is commonly associated with reptile and amphibian hosts. Additional studies are needed to further our understanding of

various *Salmonella* serovars and their association with different hosts in produce production environments, especially in other produce production environment in the US and elsewhere.

Figure 4.1 *Xba*I PFGE patterns for the representative 112 *S. enterica* isolates from environmental samples obtained from New York State (NE) and S. FL (SE) produce production environments. Band sizes (kb) are displayed at the top of the PFGE pattern images. PFGE pattern order displayed is result of BioNumerics similarity analyses using the unweighted pair group-matching algorithm (UPGMA) and the Dice correlation coefficient with a maximum space tolerance of 1.5%. ID represents the isolate designation, serovar represents confirmed serovar by traditional or molecular methods, Cornell *Xba*I # represents the PFGE type assigned by comparison of PFGE patterns to the Food Safety Laboratory (Cornell University, Ithaca, NY) database of 6,000 *Salmonella* isolates, dataset represents the study origin for the isolate, region represents New York (NYS) or S. FL (S. FL), and source represents the type of sample for the isolate.



ID	Serotype	Cornell Xba1 #	Dataset	Region	Source
FSL S10-1005	Rubislaw	NYCU.JAAX01.1201	III	SE	Water
FSL S10-1044	Rubislaw	NYCU.JAAX01.1201	III	SE	Water
FSL S10-1045	Rubislaw	NYCU.JAAX01.1201	III	SE	Water
FSL S10-0947	Braenderup	NYCU.JAAX01.1196	III	SE	Soil
FSL S10-1755	Newport	NYCU.JAAX01.0126	II	NE	Drag swab
FSL S10-1031	IIIb 50:r:z	NYCU.JAAX01.1211	III	SE	Water
FSL S10-1123	IIIb 50:r:z	NYCU.JAAX01.1211	III	SE	Water
FSL S10-1160	IIIb 50:r:z	NYCU.JAAX01.1211	III	SE	Water
FSL S10-0730	IV 40:z4,z32:-	NYCU.JAAX01.1219	I	NE	Fecal
FSL S10-0876	IV 40:z4,z32:-	NYCU.JAAX01.1219	I	NE	Water
FSL S10-1821	IV 40:z4,z32:-	NYCU.JAAX01.1219	II	NE	Water
FSL S10-1822	IV 40:z4,z32:-	NYCU.JAAX01.1219	II	NE	Soil
FSL S10-0727	IV 40:z4,z32:-	NYCU.JAAX01.1219	V	NE	Drag swab
FSL S10-1754	Senftenberg	NYCU.JAAX01.1005	II	NE	Soil
FSL S10-0951	Litchfield	NYCU.JAAX01.1197	III	SE	Soil
FSL S10-1048	Gaminara	NYCU.JAAX01.1202	III	SE	Water
FSL S10-1652	Give	NYCU.JAAX01.1217	IV	NE	Water
FSL S10-1765	Give	NYCU.JAAX01.1216	II	NE	Water
FSL S10-1814	Give	NYCU.JAAX01.1216	II	NE	Soil
FSL S10-1421	Give	NYCU.JAAX01.1215	I	NE	Water
FSL S10-1649	Give	NYCU.JAAX01.1218	IV	NE	Water
FSL S10-0560	Give	NYCU.JAAX01.1216	I	NE	Water
FSL S10-1743	Newport	NYCU.JAAX01.0121	II	NE	Soil
FSL S10-1745	Newport	NYCU.JAAX01.0121	II	NE	Drag swab
FSL S10-1747	Newport	NYCU.JAAX01.0121	II	NE	Water
FSL S10-1748	Newport	NYCU.JAAX01.0121	II	NE	Soil
FSL S10-1752	Newport	NYCU.JAAX01.0121	II	NE	Soil
FSL S10-1813	Newport	NYCU.JAAX01.0121	II	NE	Soil
FSL S10-1020	Newport	NYCU.JAAX01.1213	III	SE	Water
FSL S10-1544	Newport	NYCU.JAAX01.1212	I	NE	Water
FSL S10-1630	Newport	NYCU.JAAX01.1212	IV	NE	Water
FSL S10-1632	Newport	NYCU.JAAX01.1212	IV	NE	Water
FSL S10-1640	Newport	NYCU.JAAX01.1212	IV	NE	Water
FSL S10-1642	Newport	NYCU.JAAX01.1212	IV	NE	Water
FSL S10-1060	Newport	NYCU.JAAX01.1222	III	SE	Soil
FSL S10-1113	Newport	NYCU.JAAX01.1222	III	SE	Water
FSL S10-1115	Newport	NYCU.JAAX01.1222	III	SE	Soil
FSL S10-0955	Litchfield	NYCU.JAAX01.1198	III	SE	Drag swab
FSL S10-0985	Newport	NYCU.JAAX01.1223	III	SE	Soil
FSL S10-1560	Newport	NYCU.JAAX01.0296	I	NE	Soil
FSL S10-1570	Newport	NYCU.JAAX01.0296	I	NE	Soil
FSL S10-1819	Newport	NYCU.JAAX01.0296	II	NE	Water
FSL S10-1820	Newport	NYCU.JAAX01.0296	II	NE	Water
FSL S10-1634	Newport	NYCU.JAAX01.0296	IV	NE	Water
FSL S10-1638	Newport	NYCU.JAAX01.0296	IV	NE	Drag swab
FSL S10-0721	Newport	NYCU.JAAX01.0296	V	NE	Fecal
FSL S10-1403	Newport	NYCU.JAAX01.1221	I	NE	Soil
FSL S10-1406	Newport	NYCU.JAAX01.1221	I	NE	Fecal
FSL S10-1409	Newport	NYCU.JAAX01.1221	I	NE	Fecal
FSL S10-1228	Newport	NYCU.JAAX01.1221	III	SE	Water
FSL S10-1233	Newport	NYCU.JAAX01.1221	III	SE	Water
FSL S10-1647	Infantis	NYCU.JAAX01.1203	IV	NE	Drag swab
FSL S10-1669	Infantis	NYCU.JAAX01.1203	IV	NE	Water
FSL S10-1621	Enteritidis	NYCU.JAAX01.1225	IV	NE	Water
FSL S10-1623	Enteritidis	NYCU.JAAX01.1225	IV	NE	Soil
FSL S10-1644	Enteritidis	NYCU.JAAX01.1225	IV	NE	Drag swab
FSL S10-1646	Enteritidis	NYCU.JAAX01.1225	IV	NE	Soil
FSL S10-1081	Saphra	NYCU.JAAX01.1194	III	SE	Soil
FSL S10-1088	Saphra	NYCU.JAAX01.1194	III	SE	Soil
FSL S10-1100	Saphra	NYCU.JAAX01.1194	III	SE	Drag swab
FSL S10-1108	Saphra	NYCU.JAAX01.1194	III	SE	Water
FSL S10-1144	Saphra	NYCU.JAAX01.1194	III	SE	Water
FSL S10-1156	Saphra	NYCU.JAAX01.1194	III	SE	Water
FSL S10-1164	Saphra	NYCU.JAAX01.1194	III	SE	Water
FSL S10-1626	IIIb 50:r:z	NYCU.JAAX01.1210	IV	NE	Water
FSL S10-1628	IIIb 50:r:z	NYCU.JAAX01.1210	IV	NE	Water
FSL S10-1664	IIIb 50:r:z	NYCU.JAAX01.1210	IV	NE	Water
FSL S10-1134	Typhimurium	NYCU.JAAX01.0072	III	SE	Water
FSL S10-1568	Thompson	NYCU.JAAX01.0157	I	NE	Drag swab
FSL S10-1574	Thompson	NYCU.JAAX01.0157	I	NE	Soil
FSL S10-0354	Thompson	NYCU.JAAX01.0157	I	NE	Water
FSL S10-0860	Thompson	NYCU.JAAX01.0157	I	NE	Fecal
FSL S10-1763	Thompson	NYCU.JAAX01.0157	II	NE	Water
FSL S10-1811	Thompson	NYCU.JAAX01.0157	II	NE	Soil
FSL S10-1816	Thompson	NYCU.JAAX01.0157	II	NE	Soil
FSL S10-1817	Thompson	NYCU.JAAX01.0157	II	NE	Water
FSL S10-1818	Thompson	NYCU.JAAX01.0157	II	NE	Drag swab
FSL S10-1034	Thompson	NYCU.JAAX01.0157	III	SE	Water
FSL S10-1636	Thompson	NYCU.JAAX01.0157	IV	NE	Drag swab
FSL S10-0715	Thompson	NYCU.JAAX01.0157	V	NE	Water
FSL S10-0717	Thompson	NYCU.JAAX01.0157	V	NE	Fecal
FSL S10-1548	Thompson	NYCU.JAAX01.1199	I	NE	Water
FSL S10-1014	Thompson	NYCU.JAAX01.1199	III	SE	Drag swab
FSL S10-1131	Thompson	NYCU.JAAX01.1200	III	SE	Drag swab
FSL S10-1558	I 6,8:l:-	NYCU.JAAX01.0096	I	NE	Water
FSL S10-1663	Rubislaw	NYCU.JAAX01.1220	IV	NE	Water
FSL S10-1750	Agona	NYCU.JAAX01.1131	II	NE	Soil
FSL S10-1759	Agona	NYCU.JAAX01.1131	II	NE	Drag swab
FSL S10-1761	Agona	NYCU.JAAX01.1131	II	NE	Water
FSL S10-0975	Baildon	NYCU.JAAX01.0345	III	SE	Drag swab
FSL S10-0992	Baildon	NYCU.JAAX01.0345	III	SE	Water
FSL S10-1269	Typhimurium	NYCU.JAAX01.1208	I	NE	Water
FSL S10-1237	I 4,5,12:i:-	NYCU.JAAX01.1209	III	SE	Water
FSL S10-1241	I 4,5,12:i:-	NYCU.JAAX01.1209	III	SE	Water
FSL S10-1757	Tennessee	NYCU.JAAX01.1214	II	NE	Drag swab
FSL S10-1766	Typhimurium	NYCU.JAAX01.1207	II	NE	Water
FSL S10-1244	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-1253	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-1265	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-1411	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-0228	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-0233	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-0545	Cerro	NYCU.JAAX01.0213	I	NE	Drag swab
FSL S10-0550	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-0555	Cerro	NYCU.JAAX01.0213	I	NE	Water
FSL S10-0861	Cerro	NYCU.JAAX01.0213	I	NE	Soil
FSL S10-1741	Cerro	NYCU.JAAX01.0213	II	NE	Soil
FSL S10-1742	Cerro	NYCU.JAAX01.0213	II	NE	Soil
FSL S10-1809	Cerro	NYCU.JAAX01.0213	II	NE	Soil
FSL S10-1823	Cerro	NYCU.JAAX01.0213	II	NE	Drag swab
FSL S10-1824	Cerro	NYCU.JAAX01.0213	II	NE	Soil
FSL S10-0732	Cerro	NYCU.JAAX01.0213	V	NE	Drag swab

***Salmonella* PFGE-types show significant differences between regions.** To explore the regional *Salmonella* diversity associated with the produce production environment, we further subtyped the *Salmonella* isolates from NYS (n=80) and S. FL (n=32). We identified 37 unique *Xba*I-PFGE-types among the 112 *Salmonella* isolates across both regions (Table 2). Of the 37 PFGE-types, 23 and 17 were observed in each the NYS and S. FL regions, respectively, and three PFGE types were shared between the two regions (Table 2, Fig 1). A high level of PFGE-type diversity was observed amongst all isolates (D=0.90±0.02); in addition to within each region (D=0.92±0.03 and 0.93±0.05 for NYS and S. FL, respectively). The largest number of PFGE-types was observed among *S.* Newport isolates. The 29 *S.* Newport isolates across both regions produced eight unique PFGE-types. Of the eight *S.* Newport PFGE-types, 4 and 3 were observed exclusively in the NYS and S. FL regions, respectively, and one PFGE-type was shared between the two regions (Table 2). The isolation of unique *S.* Newport subtypes in each region may suggest certain strains are better adapted to specific ecological niches. For instance, the same strain of *S.* Newport has been repeatedly isolated over a ten year span from the eastern shore of Virginia (US) (58). This strain of *S.* Newport has been associated with at least two known outbreaks, linked to tomatoes harvested in Virginia (58), in addition to being isolated from several waterfowl and non-water fowl fecal samples collected from the eastern shore of Virginia (70). Additionally, in our study, different PFGE-types of *S. enterica* subsp. *diarizonae* 50:r:z were identified from the NYS and S. FL regions (Fig. 1). These data may indicate carriage of *Salmonella* strains by reptile or amphibian populations exclusive to NYS and S. FL. One study conducted in Mississippi (US) observed a strong association between patients infected with *S.* Javiana and contact with amphibian species endemic to the southeastern US (71). Gorski *et al.* (13) also found several *Salmonella* strains, all with the same PFGE-type, repeatedly isolated from cold blooded vertebrates and surface water in the same region in California. These findings,

along with our study data demonstrate some *Salmonella* subtypes may be more prevalent in certain areas possibly due to persistence in animal and human host populations and/or environmental pressures specific to those regions.

Only three-PFGE types were shared between the two regions, *S.* Thompson NYCU.JAAX01.0157 and NYCU.JAAX01.1199; in addition to *S.* Newport NYCU.JAAX01.1221 (Table 2). A 1,000,000 iteration simulation calculated the probability of three or fewer shared PFGE types being observed by chance between the two regions was <0.000001. This finding suggests a distinct difference between the *Salmonella* isolates recovered in the NYS and S. FL regions. *Salmonella* isolates from NYS were obtained over a three year period (2010-2012) and represented a number of produce production environments across the state (S1). Whereas *Salmonella* isolates from S. FL were obtained at one time point (Dec 2010) and represented a small subset of produce production environments (S1). Therefore, our study findings may only be applicable to the two sampled regions (e.g., findings may not be applicable to other produce production environments in US and elsewhere). For example, a study (24) conducted in North Florida (Upper Suwannee River Basin) isolated several different serovars from *Salmonella* positive water samples, as compared to serovars found among the S. FL *Salmonella* isolates reported here. In addition, our findings may not be comparable to other time periods during the year, as other studies have observed meteorological events (22, 25, 72) or field management practices (26, 73, 74) to influence pathogen prevalence in produce production environments. On the other hand, previous studies (22-24, 39, 72) in regions within the US and Canada have isolated noticeably different serovars of *Salmonella* in each study regions, further supporting *Salmonella* subtypes may be dependent on region (i.e., sampled area). For instance, *S. enterica* subsp. *arizonae* was identified in 40% of the *Salmonella* isolated from a watershed located in Georgia (US) (39) and *S.* Rubislaw was isolated most frequently from water samples

obtained from a Canadian watershed (22). Furthermore, a study investigating the distribution and diversity of *Salmonella* in a produce growing region in California primarily isolated *S.* Give (72). Our study may have some limitations (as discussed), but our data further support that *Salmonella* subtypes may differ considerably between geographical regions, and specifically subtypes differ between the sampled produce production areas in NYS and S. FL.

To summarize, our data are consistent with previous findings (17, 20, 21) that demonstrated the application of PFGE-based source tracking of foodborne pathogens in the environment. Additionally, the association between specific serovars or PFGE-types and a region may aid in pinpointing a contamination source. Lastly, this study reported here provides further evidence that food safety in the produce production environment may be best tailored to specific farms, as each farm has different characteristics, such as varying landscapes, field practices, climate, and wildlife populations.

.

REFERENCES

1. **Grimont PAD, Weill FX.** 2007. Antigenic Formulae of the *Salmonella* Serovars, 9th ed. World Health Organization Collaborating Centre for Reference and Research on Salmonella, Pasteur Institute, Paris.

2. **Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM.** 2011. Foodborne Illness Acquired in the United States-Major Pathogens. Emerg. Infect. Dis. **17:**7-15.

3. **Frenzen PD, Riggs TL, Buzby JC, Breuer T, Roberts T, Voetsch D, Reddy S.** 1999. *Salmonella* cost estimate updated using FoodNet data. Food Rev. **22:**10-15.

4. **Hanning IB, Nutt JD, Ricke SC.** 2009. Salmonellosis Outbreaks in the United States Due to Fresh Produce: Sources and Potential Intervention Measures. Foodborne Pathog. and Dis. **6:**635-648.

5. **Sivapalasingam S, Friedman CR, Cohen L, Tauxe RV.** 2004. Fresh produce: A growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. J. Food Prot. **67:**2342-2353.

6. **Painter JA, Hoekstra RM, Ayers T, Tauxe RV, Braden CR, Angulo FJ, Griffin PM.** 2013. Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998-2008. Emerg. Infect. Dis. **19:**407-415.

7. **Gould LH, Nisler AL, Herman KM, Cole DJ, Williams IT, Mahon BE, Griffin PM, Hall AJ.** 2011. Surveillance for Foodborne Disease Outbreaks - United States, 2008. MMWR **60:**1197-1202.

8. **Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV, Force CDCPT.** 2001. PulseNet: The molecular subtyping network for foodborne bacterial disease surveillance, United States. Emerg. Infect. Dis. **7:**382-389.

9. **Ranieri ML, Shi CL, Switt AIM, den Bakker HC, Wiedmanna M.** 2013. Comparison of Typing Methods with a New Procedure Based on Sequence Characterization for Salmonella Serovar Prediction. J. Clin. Microbiol. **51:**1786-1797.

10. **Uzzau S, Brown DJ, Wallis T, Rubino S, Leori G, Bernard S, Casadesus J, Platt DJ, Olsen JE.** 2000. Host adapted serotypes of *Salmonella* enterica. Epidemiol. Infect. **125:**229-255.

11. **Smith HW, Jones JET.** 1967. Observations on experimental oral infection with *Salmonella* dublin in calves and *Salmonella* choleraesuis in pigs. J. Pathol. Bacteriol. **93:**141-156.

12. **de Sa IVA, Solari CA.** 2001. *Salmonella* in Brazilian and imported pet reptiles. Braz. J. Microbiol. **32:**293-297.

13. **Gorski L, Jay-Russell MT, Liang AS, Walker S, Bengson Y, Govoni J, Mandrell RE.** 2013. Diversity of Pulsed-Field Gel Electrophoresis Pulsotypes, Serovars, and Antibiotic Resistance Among *Salmonella* Isolates from Wild Amphibians and Reptiles in the California Central Coast. Foodborne Pathog. Dis. **10:**540-548.

14. **Schroter M, Roggentin P, Hofmann J, Speicher A, Laufs R, Mack D.** 2004. Pet snakes as a reservoir for *Salmonella* enterica subsp diarizonae (serogroup IIIb): a prospective study. Appl. Environ. Microbiol. **70:**613-615.

15. **Galanis E, Wong D, Patrick ME, Binsztein N, Cieslik A, Chalermchaikit T, Aidara-Kane A, Ellis A, Angulo FJ, Wegener HC, Surv WHOGS.** 2006. Web-based surveillance and global *Salmonella* distribution, 2000-2002. Emerg. Infect. Dis. **12:**381-388.

16. **Lawson B, Hughes LA, Peters T, de Pinna E, John SK, Macgregor SK, Cunningham AA.** 2011. Pulsed-Field Gel Electrophoresis Supports the Presence of Host-Adapted *Salmonella enterica* subsp *enterica* Serovar Typhimurium Strains in the British Garden Bird Population. Appl. Environ. Microbiol. **77:**8139-8144.

17. **Cooley M, Carychao D, Crawford-Miksza L, Jay MT, Myers C, Rose C, Keys C, Farrar J, Mandrell RE.** 2007. Incidence and Tracking of *Escherichia coli* O157:H7 in a Major Produce Production Region in California. Plos One **2:**16.

18. **Uesugi AR, Danyluk MD, Mandrell RE, Harris LJ.** 2007. Isolation of *Salmonella* Enteritidis phage type 30 from a single almond orchard over a 5-year period. J. Food Prot. **70:**1784-1789.

19. **Patchanee P, Molla B, White N, Line DE, Gebreyes WA.** 2010. Tracking *Salmonella* Contamination in Various Watersheds and Phenotypic and Genotypic Diversity. Foodborne Pathog. Dis. **7:**1113-1120.

20. **Ram JL, Ritchie RP, Fang JW, Gonzales FS, Selegean JP.** 2004. Sequence-based source tracking of *Escherichia coli* based on genetic diversity of beta-glucuronidase. J. Environ. Qual. **33:**1024-1032.

21. **Furukawa T, Yoshida T, Suzuki Y.** 2011. Application of PFGE to source tracking of faecal pollution in coastal recreation area: a case study in Aoshima Beach, Japan. J. Appl. Microbiol. **110:**688-696.

22. **Jokinen C, Edge TA, Ho S, Koning W, Laing C, Mauro W, Medeiros D, Miller J, Robertson W, Taboada E, Thomas JE, Topp E, Ziebell K, Gannon VPJ.** 2011. Molecular subtypes of *Campylobacter* spp., *Salmonella enterica*, and *Escherichia coli* O157:H7 isolated from faecal and surface water samples in the Oldman River watershed, Alberta, Canada. Water Res. **45:**1247-1257.

23. **Walters SP, Gonzalez-Escalona N, Son I, Melka DC, Sassoubre LM, Boehm AB.** 2013. *Salmonella enterica* Diversity in Central Californian Coastal Waterways. Appl. Environ. Microbiol. **79:**4199-4209.

24. **Rajabi M, Jones M, Hubbard M, Rodrick G, Wright AC.** 2011. Distribution and genetic diversity of *Salmonella enterica* in the Upper Suwannee River. Int. J. Microbiol. **2011:**Article ID 461321.

25. **Strawn LK, Fortes ED, Bihn EA, Nightingale KK, Grohn YT, Worobo RW, Wiedmann M, Bergholz PW.** 2013. Landscape and meteorological factors affecting prevalence of three food-borne pathogens in fruit and vegetable farms. Appl. Environ. Microbiol. **79:**588-600.

26. **Strawn LK, Grohn YT, Warchocki S, Worobo RW, Bihn EA, Wiedmann M.** 2013. Risk Factors Associated with *Salmonella* and *Listeria monocytogenes* Contamination of Produce Fields. Appl. Environ. Microbiol.**79:**7618-7627.

27. **Andrews WH, Hammack T** 2011, posting date. BAM: Detection and Enumeration of *Salmonella*, Ch. 5. Food and Drug Administration.

28. **Kim JS, Lee GG, Park JS, Jung YH, Kwak HS, Kim SB, Nam YS, Kwon ST.** 2007. A novel multiplex PCR assay for rapid and simultaneous detection of five pathogenic bacteria: *Escherichia coli* O157 : H7, *Salmonella*, *Staphylococcus aureus*, *Listeria monocytogenes*, and *Vibrio parahaemolyticus*. J. Food Prot. **70:**1656-1662.

29. **Michael GB, Simoneti R, da Costa M, Cardoso M.** 2003. Comparison of different selective enrichment steps to isolate *Salmonella* sp from feces of finishing swine. Braz. J. Microbiol. **34:**138-142.

30. **Schonenbrucher V, Mallinson ET, Bulte M.** 2008. A comparison of standard cultural methods for the detection of foodborne *Salmonella* species including three new chromogenic plating media. Int. J. Food Microbiol. **123:**61-66.

31. **Ribot EM, Fair MA, Gautom R, Cameron DN, Hunter SB, Swaminathan B, Barrett TJ.** 2006. Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157 : H7 *Salmonella*, and *Shigella* for PulseNet. Foodborne Pathog. Dis. **3:**59-67.

32. **Hunter SB, Vauterin P, Lambert-Fair MA, Van Duyne MS, Kubota K, Graves L, Wrigley D, Barrett T, Ribot E.** 2005. Establishment of a universal size standard strain for use with the PulseNet standardized pulsed-field gel electrophoresis protocols:

Converting the national databases to the new size standard. J. Clin. Microbiol. **43:**1045-1050.

33. **Herrera-Leon S, Ramiro R, Arroyo M, Diez R, Usera MA, Echeita MA.** 2007. Blind comparison of traditional serotyping with three multiplex PCRs for the identification of *Salmonella* serotypes. Res. Microbiol. **158:**122-127.

34. **Imre A, Olasz F, Nagy B.** 2005. Development of a PCR system for the characterisation of *Salmonella* flagellin genes. Acta Vet. Hung. **53:**163-172.

35. **Mortimer CKB, Peters TM, Gharbia SE, Logan JMJ, Arnold C.** 2004. Towards the development of a DNA-sequence based approach to serotyping of *Salmonella enterica*. BMC Microbiol. **4:**10.

36. **R-Team** 2004, posting date. R: a language and environment for statistical computing. The R Foundation for Statistical Computing.

37. **Simpson EH.** 1949. Measurement of Diversity. Nat. **163:**688-688.

38. **Grundmann H, Hori S, Tanner G.** 2001. Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms. J. Clin. Microbiol. **39:**4190-4192.

39. **Haley BJ, Cole DJ, Lipp EK.** 2009. Distribution, Diversity, and Seasonality of Waterborne *Salmonellae* in a Rural Watershed. Appl. Environ. Microbiol. **75:**1248-1255.

40. **Corrente M, Madio A, Friedrich KG, Greco G, Desario C, Tagliabue S, D'Incau M, Campolo M, Buonavoglia C.** 2004. Isolation of *Salmonella* strains from reptile feces and comparison of different culture media. J. Appl. Microbiol. **96:**709-715.

41. **Rall VLM, Rall R, Aragon LC, da Silva MG.** 2005. Evaluation of three enrichment broths and five plating media for *Salmonella* detection in poultry. Braz. J. Microbiol. **36:**147-150.

42. **Vassiliadis P.** 1983. The Rappaport-Vassiliadis (RV) enrichment medium for the isolation of *Salmonellas* - An overview. J. Appl. Bacteriol. **54:**69-74.

43. **Papavassilio J, Samaraki-Lyberopoulou V, Piperakis G.** 1969. Production of tetrathionate reductase by *Salmonella*. Can. J. Micrbiol. **15:**238-240.

44. **Mallinson ET, Miller RG, de Rezende CE, Ferris KE, deGraft-Hanson J, Joseph SW.** 2000. Improved plating media for the detection of *Salmonella* species with typical and atypical hygrogen sulfide production. J. Vet. Diagn. Invest. **12:**83-87.

45. **McDonough PL, Shin SJ, Lein DH.** 2000. Diagnostic and public health dilemma of lactose-fermenting *Salmonella enterica* serotype Typhimurium in cattle in the Northeastern United States. J. Clin. Microbiol. 38:1221-1226.

46. **Puhr N, O'Hara C, Miller M, Bopp C, Wells J.** 1994. Evaluation of methods for isolation and identification of lactose-fermenting *Salmonella*. Abstr. Gen. Meet. Am. Soc. Microbiol. **94:**557.

47. **Tyc Z, Szych J, Kaluzewski S.** 1989. Selection of methods of the detection of lactose-fermenting *Salmonella* strains and evaluation of their usefulness in diagnostic studies. Med. Dosw. Mikrobiol. **41:**106-114.

48. **Centers for Disease Control and Prevention.** 2008. *Salmonella* Surveillance: Annual Summary, 2006. US Department of Health and Human Services.

49. **Callaway TR, Edrington TS, Anderson RC, Byrd JA, Nisbet DJ.** 2008. Gastrointestinal microbial ecology and the safety of our food supply as related to *Salmonella*. J. Anim. Sci. **86:**E163-172.

50. **Jackson BR, Griffin PM, Cole D, Walsh KA, Chai SJ.** 2013. Outbreak-associated *Salmonella enterica* serotypes and food Commodities, United States, 1998-2008. Emerg. Infect. Dis. **19:**1239-1244.

51. **Centers for Disease Control and Prevention.** 2003. Reptile-associated salmonellosis--selected states, 1998-2002. MMWR **52:**1206-1209.

52. **Centers for Disease Control and Prevention.** 2010. Multistate outbreak of human *Salmonella* Typhimurium infections associated with aquatic frogs - United States, 2009. MMWR **58:**1433-1436.

53. **Ackman DM, Drabkin P, Birkhead G, Cieslak P.** 1995. Reptile associated Salmonellosis in New York State. J. Pediatr. Infect. Dis. **14:**955-959.

54. **Hoelzer K, Switt AIM, Wiedmann M.** 2011. Animal contact as a source of human non-typhoidal salmonellosis. Vet. Res. **42:**28.

55. **Hendriksen RS, Vieira AR, Karlsmose S, Wong D, Jensen AB, Wegener HC, Aarestrup FM.** 2011. Global Monitoring of *Salmonella* Serovar Distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank: Results of Quality Assured Laboratories from 2001 to 2007. Foodborne Pathog. Dis. **8:**887-900.

56. **Sangal V, Harbottle H, Mazzoni CJ, Helmuth R, Guerra B, Didelot X, Paglietti B, Rabsch W, Brisse S, Weill FX, Roumagnac P, Achtman M.** 2010. Evolution and

Population Structure of *Salmonella enterica* Serovar Newport. J. Bacteriol. **192:**6465-6476.

57. **Food and Drug Administration.** 2013. Environmental Assessment: Factors Potentially Contributing to the Contamination of Fresh Whole Cantaloupe Implicated in a Multi-State Outbreak of Salmonellosis FDA. http://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm341476.htm.

58. **Greene SK, Daly ER, Talbot EA, Demma LJ, Holzbauer S, Patel NJ, Hill TA, Walderhaug MO, Hoekstra RM, Lynch MF, Painter JA.** 2008. Recurrent multistate outbreak of *Salmonella* Newport associated with tomatoes from contaminated fields, 2005. Epidemiol. Infect. **136:**157-165.

59. **Harris LJ, Farber JN, Beuchat LR, Parish ME, Suslow TV, Garrett EH, Busta FF.** 2003. Outbreaks Associated with Fresh Produce: Incidence, Growth, and Survival of Pathogens in Fresh and Fresh Cut Produce Comprehensive Rev. Food Sci. Food Safety **2:**78-141.

60. **Van Kessel JAS, Karns JS, Wolfgang DR, Hovingh E.** 2013. Regional Distribution of Two Dairy-Associated *Salmonella enterica* Serotypes. Foodborne Pathog. Dis. **10:**448-452.

61. **Cummings KJ, Warnick LD, Elton M, Rodriguez-Rivera LD, Siler JD, Wright EM, Groehn YT, Wiedmann M.** 2010. *Salmonella enterica* Serotype Cerro Among Dairy Cattle in New York: An Emerging Pathogen? Foodborne Pathog. Dis. **7:**659-665.

62. **Hoelzer K, Cummings KJ, Wright EM, Rodriguez-Rivera LD, Roof SE, Switt AIM, Dumas N, Root T, Schoonmaker-Bopp DJ, Grohn YT, Slier JD, Warnick LD,**

Hancock DD, Davis MA, Wiedmann M. 2011. *Salmonella* Cerro isolated over the past twenty years from various sources in the US represent a single predominant pulsed-field gel electrophoresis type. Vet. Microbiol. **150:**389-393.

63. **Mammina C, Cannova L, CarfiPavia S, Nastasi A.** 2000. Endemic presence of *Salmonella enterica* serotype Cerro in southern Italy. Euro. Surveill. **5:**84-86.

64. **Dondero NC, Thomas CT, Khare M, Timoney JF, Fukui GM.** 1977. *Salmonella* in Surface waters of central New York State. Appl. Environ. Microbiol. **33:**791-801.

65. **Campbell JV, Mohle-Boetani J, Reporter R, Abbott S, Farrar J, Brandl M, Mandrell R, Werner SB.** 2001. An outbreak of *Salmonella* serotype Thompson associated with fresh cilantro. J. Infect. Dis. **183:**984-987.

66. **Nygard K, Lassen J, Vold L, Andersson Y, Fisher I, Lofdahl S, Threlfall J, Luzzi I, Peters T, Hampton M, Torpdahl M, Kapperud G, Aavitsland P.** 2008. Outbreak of *Salmonella* Thompson infections linked to imported rucola lettuce. Foodborne Pathog. Dis. **5:**165-173.

67. **Benassi FO, Martinez Vazquez F, Eiguer T, Bendersky S, Martos MA.** 1985. Isolation of New Serovars of Salmonella from Steams. Rev. Argentina Microbiol. **17:**149-156.

68. **Souza LCD, Laria ST, Paim GV.** 1992. *Salmonellas* and fecal coliforms in animal drinking water. Rev. Saude Publica **26:**321-327.

69. **Mohle-Boetani JC, Reporter R, Werner SB, Abbott S, Farrar J, Waterman SH, Vugia DJ.** 1999. An outbreak of *Salmonella* serogroup saphra due to cantaloupes from Mexico. J. Infect. Dis. **180:**1361-1364.

70. **Gruszynski K, Pao S, Kim C, Toney D, Wright K, Ross PG, Colon A, Levine S.** 2013. Evaluating wildlife as a potential source of *Salmonella* serotype Newport (JJPX01.0061) contamination of tomatoes on the eastern shore of Virginia. Zoonoses Public Hlth. doi:10.1111/zph.12061 [Epub ahead of print].

71. **Srikantiah P, Lay JC, Hand S, Crump JA, Campbell J, Van Duyne MS, Bishop R, Middendor R, Currier M, Mead PS, Molbak K.** 2004. *Salmonella enterica* serotype Javiana infections associated with amphibian contact, Mississippi, 2001. Epidemiol. Infect. **132:**273-281.

72. **Gorski L, Parker CT, Liang A, Cooley MB, Jay-Russell MT, Gordus AG, Atwill ER, Mandrell RE.** 2011. Prevalence, Distribution, and Diversity of *Salmonella enterica* in a Major Produce Region of California. Appl. Environ. Microbiol. **77:**2734-2748.

73. **Mukhejee A, Speh D, Diez-Gonzalez F.** 2007. Association of farm management practices with risk of *Escherichia coli* contamination in pre-harvest produce grown in Minnesota and Wisconsin. Intern. J. Food Micrbiol. **120:**296-302.

74. **Park S, Navratil S, Gregory A, Bauer A, Srinath I, Jun M, Szonyi B, Nightingale K, Anciso J, Ivanek R.** 2013. Generic *Escherichia coli* contamination of spinach at the preharvest stage: Effects of farm management and environmental factors. Appl. Environ. Microbiol. **79:**4347-4358.

CHAPTER 5

CONCLUSIONS

As fresh fruits and vegetables continue to be a vehicle for foodborne disease, produce growers are expected to be vigilant in their efforts to minimize the risk of produce contamination. In 2015, produce growers will be expected to begin compliance with the Food and Drug Administration (FDA) Food Safety Modernization Act (FSMA), Rule for Produce Safety, which will set science-based standards for (i) agricultural water, (ii) biological soil amendments of animal origin, (iii) health and hygiene, (iv) animals in the growing area, and (v) equipment, tools and buildings. In anticipation of these regulatory changes, my research aimed to identify sources and vectors of preharvest contamination supplying growers and food safety professionals with science-based recommendations and tools for reducing the likelihood of produce contamination in the preharvest environment.

To accomplish this, the overarching goal of my research was to better understand the ecology of *Salmonella* and *L. monocytogenes* in the preharvest environment. We hypothesized that pathogen prevalence and persistence was likely dependent on the farm landscape and that certain management practices interact with that landscape to influence the risk of pathogen contamination in the production environment and in the produce. As a result, each component of my research utilized a scientific application to model the association between pathogen ecology and risk of preharvest contamination, including Geographic Information Systems (GIS)-enabled modeling, risk management, and subtyping-based source tracking. Collectively, these scientific applications allowed us to explore potential reservoirs and transmission pathways of pathogen contamination in the production environment.

In the first study, we identified a number of key geographical factors that influenced the prevalence of *Salmonella* and *L. monocytogenes* in the produce preharvest environment by

employing GIS-enabled modeling. Several of these factors used to model pathogen prevalence and persistence in the environment by GIS software were remotely-sensed and publically available. A series of rules were determined (based on factors) to predict the pathogen prevalence for a spatial location, thereby classifying areas on farms as favorable or unfavorable reservoirs for pathogens. Our results determined drainage class and soil moisture (e.g., available water storage) were the most important factors when predicting a favorable *Salmonella* reservoir. For instance, field locations with poorly drained soil were predicted to have a *Salmonella* prevalence of 9%, while field locations with moderately well drained soil had a predicted *Salmonella* prevalence of only 1%. On the other hand, our results determined proximity to pastures (e.g., domestic animal operations, grazing lands), impervious surfaces (e.g., roadways, urban development), water (e.g., pond, creek), and level of soil moisture were the most important factors when predicting a favorable *L. monocytogenes* reservoir. For example, field locations near an impervious surface were predicted to have an *L. monocytogenes* prevalence of 20%, while field locations away from impervious surfaces had a predicted *L. monocytogenes* prevalence of only 5%. Recent weather events, measurable precipitation and below average temperatures 3 days antecedent to environmental sample collection where shown to increase the likelihood of a *Salmonella* and *L. monocytogenes* positive sample, respectively. These findings demonstrated that proximity to various land-uses, soil properties and climate played a major role in *Salmonella* and *L. monocytogenes* ecology in produce fields, and that under certain conditions may greatly increase the risk of produce contamination.

Thus, with the development of these factor-derived predicted prevalence data, we can use GIS software to test predictions on the spatial locations of *Salmonella* and *L. monocytogenes* environmental reservoirs on commercial produce farms. Future research includes validation of GIS algorithms to identify areas predicted to be reservoirs of pathogens. The application of GIS-

enabled modeling to the produce industry is very realistic. GIS-based technologies are becoming more common in farming and many produce growers are already using GIS tools, such as precision agriculture, to optimize fertilizer or pesticide application to a field.

In the second study, we determined field-level management practices that were associated with a *Salmonella*- and *L. monocytogenes*-positive field by using questionnaires and environmental sampling, in concert with traditional statistical methods. This study provided quantitative data on management practices that represented risk factors for produce field contamination, which would allow growers to implement changes to field practices or develop preventative controls to reduce the risk of preharvest contamination. Multivariable analysis showed manure application and presence of a buffer zone were significant risk factors influencing the likelihood of *Salmonella* isolation in fields. Specifically, detection of *Salmonella* in fields was more likely if manure was applied to fields within a year prior of sample collection, whereas detection of *Salmonella* in fields was less likely if fields had a buffer zone. On the other hand, multivariable analysis showed reported observation of wildlife, irrigation, and soil cultivation were significant risk factors influencing the likelihood of *L. monocytogenes* isolation in fields. Reported observation of wildlife and irrigation in a field within 3 days prior to sample collection each increased the odds of a *L. monocytogenes*-positive field. Soil cultivation in a field within 7 days prior to sample collection also increased the odds of a *L. monocytogenes*-positive field. These findings demonstrated that specific management practices do interact with field ecology, and under certain conditions may influence the risk of produce contamination in the preharvest environment.

The identification of these field-level management risk factors will aid in the development of prevention-based preharvest food safety recommendations for growers to implement on their farms. Some recommendations may only consist of small changes to field

practices; however, those changes may greatly reduce the risk of produce preharvest contamination. We can also test intervention strategies for factors identified to be risk factors of field contamination. For instance, a future study may further investigate field buffer zones and risk reduction by analyzing the effect of buffer zone type (e.g., bare ground or vegetation buffer zones), and width (e.g., 5 m, 10 m, 30 m) on pathogen reduction.

In the third study, we investigated the application of subtyping-based source tracking of *Salmonella* in the environment by characterizing *Salmonella* isolates from two geographically diverse produce growing regions (New York State (NYS) and South Florida (S. FL)) in the US. This study provided evidence to suggest that there are regional differences in *Salmonella* populations in production environments. We were able to quantify the probability of isolates sharing pulsed field gel electrophoresis (PFGE) types between the two regions and statistically show that *Salmonella* isolates were distributed regionally. This may demonstrate certain *Salmonella* strains have a greater prevalence and/or persistence in specific ecological niches or are associated with specific hosts prevalent in that region. Furthermore, our data demonstrated potential sources and vectors of preharvest contamination may be extrapolated from subtype data. For example, *S.* Cerro was isolated exclusively from the NYS production environment and is highly associated with dairy cattle hosts. A potential source of preharvest contamination in some NYS produce farms may be from dairy cattle production environments. This hypothesized source of contamination may be further examined by the use of GIS (previously discussed) to map proximity of dairy cattle production environments to produce farm/fields. Additionally, this source of contamination may also be connected to the application of manure of dairy cattle origin to fields.

These research findings suggest that preharvest produce safety may be best managed at a regional level, at least in the US. The NYS and S. FL regions differ quite drastically in

landscape, climate and produce growing practices; thus, it seems likely that sources and vectors of contamination would be different in production environments (e.g., *S.* Cerro specific to the NYS region, as previously discussed). However, other studies are needed to investigate regional distributions of *Salmonella* in other US producing growing regions (e.g., Southwest, Northwest, Midwest), in comparison with the results discussed here. Additionally, it may be judicious to observe how *L. monocytogenes* populations are distributed in different produce growing regions. Therefore, forthcoming studies will investigate the application of subtyping-based source tracking of *L. monocytogenes* by characterizing *L. monocytogenes* isolates obtained from diverse regions in NYS (approximately 300 representative isolates in an existing collection), and other produce growing regions in future collaborations.

These studies further our understanding of *Salmonella* and *L. monocytogenes* ecology in the environment, and specifically the production environment. We have identified several factors that influence the likelihood of preharvest contamination via different approaches to modeling pathogen contamination as an ecological process. The potential for a technology to be fabricated that ensures removal of *Salmonella* and *L. monocytogenes* from the environment is highly unlikely; therefore, the only solution is to limit the opportunities for produce to become contaminated. These studies outline science-based strategies to minimize contamination risks in the production environment by aiding growers (i) to make more informed decisions related to field practices, (ii) to have targeted pathogen surveillance programs, and (iii) to alter planting schemes based on associated risk.